The role of the kauralexin biosynthetic pathway in maize defences against the fall armyworm, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera, Noctuidae)

by

Robert Luke Jansen van Vuuren

Submitted in fulfilment of the requirements for the degree

Magister Scientiae

in the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

December 2021

Primary supervisor: Prof. Dave Berger

Co-supervisors: Prof. Kerstin Krüger

Dr. Nicky Creux

Table of Contents

List of Abbreviations

An1	Anther ear 1
An2	Anther ear 2
Bt	*Bacillus thuringiensis*
CPS	Copalyl diphosphate synthase
COI	Cytochrome oxidase I
CS	Corn strain
Exosap	Exonuclease 1 and shrimp alkaline phosphatase
GGPP	Geranyl-geranyl diphosphate
Min	Minute
mL	Milliliter
mM	Millimolar
PCR	Polymerase chain reaction
RS	Rice strain
SNP	Single nucleotide polymorphism
Tpi	Triose phosphate isomerase
TPS	Terpene synthase
µL	Microliter
µM	Micromolar

List of Figures

List of Tables

Declaration

I, Robert Luke Jansen van Vuuren, declare that the dissertation, which I hereby submit for the degree *Magister Scientiae* at the University of Pretoria, is my own and has not been previously submitted by me for a degree at this or any other tertiary institution.

Signature: _____

Date: _____

Acknowledgements

Psalm 121:1-2 “I lift up my eyes to the hills. From where does my help come from? My help comes from the Lord, who made heaven and earth.” Firstly, I want to thank my Heavenly Father for His sustaining power, grace and mercy over this period. It is by His hand that I knew peace and my strength was refreshed. I dedicate this dissertation to Him.

I would like to thank my supervisors, Prof. Dave Berger, Prof. Kerstin Krüger and Dr. Nicky Creux. It is with their help and guidance that I am able to submit this dissertation. To Prof. Berger, thank you for your insight and direction in this dissertation, for your questions which challenged me to understand more and ultimately be a develop as a researcher. To Prof. Krüger, thank you for your practical guidance in rearing the fall armyworm, the use of your fall armyworm culture and your vast entomological knowledge that was of great benefit in completing this project. To Dr. Creux, thank you for remaining on as one of my supervisors, your detailed feedback has been extremely helpful in developing my writing, thank you as well for your encouragement and willingness to motivate. I am truly grateful to have been under the guidance of each of you.

To the many people who helped me in this project, I am forever thankful. The staff at the Zoology and Entomology department, specifically Janine Snyman and Nikki Visagie. Janine, thank you for being a source of encouragement and wisdom, your love for the Lord and overflowing joy salted your words. Nikki, thank you for your help rearing the fall armyworm culture, especially during the heavy lockdown periods, without your help completing this project would have been severely delayed. Thank you to everyone which made up the research group I was privileged to be part of, the Molecular Plant-Pathogen Interactions (MPPI) group. Thank you especially to Ingrid, Trystan and Carla for your help setting up the fall armyworm herbivory experiments into the night, I am grateful for your selflessness and support. Thank you to Dr. Carel Oosthuizen, for your willingness to teach and review my protein modelling work, without your help this part of the study would not have been possible.

I also want to thank the Council for Scientific and Industrial Research (CSIR) as my bursar over the study period. I appreciate and am grateful for the financial support.

Thank you to my family, Lichelle, Ava, Leny, Mark, Karin, Taryn and Chad for every word of encouragement and prayer. Your prayer has carried me through and now we can say together that it is done. Thank you to my friends and church family, you have also been an integral part in me succeeding in completing this dissertation.

Finally, to my wife, Nicole, thank you for your love and support. You were with me in all the times of discouragement, hopelessness and anxiety. I know what I endured, you endured as well. You have been a great source of strength, wisdom and encouragement. I love you!

Preface

Maize is a staple food source for millions of people across the African continent. A threat to maize production is therefore a threat to food security. The increase spread of alien insect pests threatens agriculture in Africa, one such pest is the fall armyworm, *Spodoptera frugiperda*. The fall armyworm is a polyphagous, lepidopteran pest which has proven to be a great threat to maize production in the Western Hemisphere where it is endemic. Investigating maize defence compounds may provide insight to specific chemical compounds which are effective against fall armyworm larvae as either deterrents, antifeedants or toxins. Kauralexins are a class of maize diterpenoids which have shown to exhibit antimicrobial properties and when applied externally to maize plants, have resulted in significantly lower herbivory by European corn borer larvae. Studying kauralexins, their role in maize defences and their effect on fall armyworm larvae may aid in developing new control strategies which is critical to preventing outbreaks and major yield losses.

The **first chapter** of this dissertation is a review of maize chemical defences against biotic factors such as fungi and insect herbivory. Terpenoids, especially the chemical class, kauralexins is discussed. An overview of the fall armyworm, its biology, the different strains, recent introduction into Africa and its impact on maize production is also presented.

The **second chapter** is in the format of an issue article to be submitted to the Journal of Integrated Pest Management. This journal does not follow the conventional research article format. The second chapter focuses on the two fall armyworm strains, namely the rice and corn strains, their differences, presence in Africa and their differentiation using genetic markers within a laboratory reared culture. The importance of accurate strain identification, the pitfalls of defining strain identity using either *COI* and *Tpi* markers in isolation and suggestions how to improve strain determination is discussed. The term corn is used instead of maize for journal submission purposes.

The **third chapter** focused on the maize An2 protein and its role in the kauralexin biosynthesis pathway. A bioinformatics approach is taken in which the An2 protein is compared across

multiple maize inbred lines through amino acid sequence comparison, protein modelling and ligand docking. Binding affinity of the An2 protein to the ligand is correlated to total kauralexin synthesis data present in published literature.

The **fourth chapter** investigates the role of the kauralexin biosynthesis pathway in the maize defences against fall armyworm larvae. Non-choice and choice herbivory assays are performed to determine whether kauralexins potentially exhibit deterrent or toxic qualities to fall armyworm larvae resulting in less feeding on maize synthesising these compounds.

The **fifth chapter** contains the concluding remarks of this dissertation. In this chapter the previous chapters are concluded, the pitfalls of this study are highlighted and the future studies which may come from this study are discussed.

Chapter 1: Literature Review

The synthesis of kauralexins in maize and the Fall Armyworm in Africa

---

1.1 Introduction

Zea mays L. (Poaceae; maize) is an important crop plant, which is cultivated worldwide and forms a major proportion of global cereal production. This crop is a staple food for millions of people worldwide and is currently a major food source for over 200 million people in Africa (Day *et al.* 2017). This crop is also used as animal feed, for mainly cattle, swine and poultry, and for industrial purposes to produce starch, sweetener and ethanol (James 2003). The increase in globalisation of maize, coupled with increasing climate change has resulted in the invasion and spread of pest insects and pathogens (Bebber *et al.* 2014). The pests and pathogens, along with abiotic pressures, such as extreme drought and heat have consequently constrained maize production (Fakorede *et al.* 2003).

The development of maize varieties which are more resistant to pests and pathogens require further research and understanding of maize genetics, particularly those which underlies maize defences. Phytoalexins, a group of plant defence compounds, which are synthesized in elevated quantities in response to pest and pathogen attack, can act as toxins and/or deterrents to the invading pests and pathogens (Schmelz *et al.* 2011). A specialised group of terpenoid phytoalexins, termed kauralexins, have been identified in maize and noted to increase defences against various *Fusarium* species and deter feeding of the European corn borer (Harris *et al.* 2005, Schmelz *et al.* 2011, Vaughan *et al.* 2015). This class of compounds has not been extensively researched and may play an integral role in the maize defences against other economically important pests and pathogens.

The fall armyworm, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera, Noctuidae), originally from the Americas, has recently been introduced into Africa. The first occurrences were reported in West and Central Africa in 2016 (Goergen *et al.* 2016) and has since spread rapidly in Africa, reports of the pest in South Africa being as early as 2017 (Jacobs *et al.* 2018). The fall armyworm is a polyphagous pest, known to feed on a large variety of crops, including major economic crops such as maize, rice and sorghum (Day *et al.* 2017). In the United States, the fall armyworm has been a problem for maize production for over 100 years, with outbreaks resulting in significant reductions in yield (Luginbill 1928). The recent introduction in Africa has resulted in concern regarding the impact this pest will have on maize production and whether control strategies already available in Africa will be effective.

Understanding maize defences against pests and pathogens is critical in developing new maize varieties which could be used in integrated control strategies. The aim of this literature review is to (1) lay a foundation for the maize defences, particularly kauralexins, (2) discuss research which has already been carried out on kauralexins, and (3) discuss the significance of lepidopteran pests of maize in Africa, focusing on the fall armyworm. This literature review will also investigate the type of experiments available which could be used to investigate the defensive role of these compounds against the fall armyworm.

1.2 The History, Consumption and Development of Maize

The origin and evolution of *Z. mays* has been of interest to many different biological fields and have been researched extensively. The teosinte hypothesis, a hypothesis proposed in the late 1930s by George Beadle, is the accepted hypothesis for the genetic origin of maize (Bennetzen *et al.* 2001). This hypothesis proposes that modern maize originated from an ancient *teosinte,* a wild grass part of the genus *Zea* (Eubanks 2001, Ranum *et al.* 2014). Archaeological sites in Mexico provide evidence for the domestication of this crop between 5 000 and 10 000 years ago (Sluyter and Dominguez 2006). These discoveries have supported the notion that *Z. mays* originated in Central America in Mexico (Ranum *et al.* 2014).

Today maize is one of the most important crops in the world based on total production and consumption (Zampieri *et al.* 2019). In Africa this crop is a major food source and constitutes the staple food crop for over 200 million across the continent, with South Africa's production alone being 11,27 million tonnes in 2019 (Day *et al.* 2017, FAO 2019). A major factor which has allowed the globalisation and increasing production of maize has been the development of improved maize varieties through maize breeding. Maize breeding has enabled people to select for and produce maize varieties which are higher yielding and more resistant to pests, pathogens and environmental stressors (Crow 1998, Pratt and Gordon 2006, Carena *et al.* 2009, Bender *et al.* 2013). Maize hybrids and genetically modified (GM) maize constitutes the majority of the total maize production of the world (Ranum *et al.* 2014). Maize hybrids are developed by crossing well-adapted field varieties which have been extensively inbred to produce genetically stable, maize inbred lines (Șuteu *et al.* 2013). Inbred maize lines have also been used for research purposes to further understand maize genetics, such as the maize genetics which underlie plant defences against pests and pathogens (Lee 1994).

1.3 Maize Defences and Secondary Metabolites

Plants have unique innate defence systems which serve to increase their resistance to a variety of insect herbivores and phytopathogens, such as fungi and bacteria. Plant defence compounds are either synthesized constitutively, termed phytoanticipins, or are synthesized *de novo* in response to pathogen invasion or herbivory, termed phytoalexins (Meyer *et al.* 2016). Plant cells recognize pathogens and insect herbivory by identifying cues in the form of fragments of flagellin from bacterial cells or compounds in insect oral secretions. These cues are identified and result in activation of signal transduction pathways, namely phosphorylation cascades, calcium ion fluxes and increases in the jasmonate (JA) pathway (Howe and Jander 2008). These pathways stimulate the production of plant defence compounds, also termed secondary defence metabolites in response to the perceived pests or pathogens. Some of these secondary metabolite classes, which express defensive properties, include terpenoids, alkaloids, tannins, saponins, glucosinolates, furanocoumarins and cyanogenic glycosides (Howe and Jander 2008). Combinations of some of these different secondary metabolites have been shown to have synergistic effects on plant defences (Howe and Jander 2008). The focus of this study in on terpenoids.

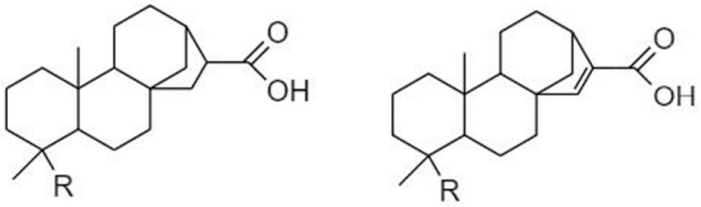
1.3.1 Maize terpenoids

Terpenoids are a metabolically diverse class of secondary metabolites, which are known to consist of more than 40 000 unique structures which play an active role in plant defence (Aharoni *et al.* 2005). A variety of non-volatile terpenoids have been known to play the role of secondary metabolite phytoalexins within the economically important crops, maize and rice (Schmelz *et al.* 2011, Yamane 2013). Maize produces a wide range of volatile and non-volatile terpenoids that are known to be synthesized as a response to biotic attack (Block *et al.* 2019). These chemical compounds are formed via two pathways, the mevalonate pathway in the cytosol and the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway in the plastids (Mahmoud

and Croteau 2002, Rodríguez-Concepción and Boronat 2002). The synthesis pathways of the terpenoid precursor molecules are further discussed in the Chapter 3.1. The prenyl diphosphate molecule, geranylgeranyl diphosphate (GGDP), is an important terpenoid precursor molecule which is utilised in the Gibberellin synthesis pathway (primary metabolite pathway) as well as the secondary metabolite pathways synthesizing dolabralexins and kauralexins.

1.3.2 Kauralexins

Kauralexins are labdane-related diterpenoids for which there are a variety of different structures (Murphy *et al.* 2018). These form two groups of compounds, namely kauralexins A and B. Kauralexin A's consist of *ent*-kaurran-17-oic acid (A1), *ent*-kauran-17,19-dioic acid (A2), *ent*-kaur-19-al-17-oic acid (A3) and *ent*-kaur-19-ol-17-oic acid (A4). Kauralexin B's consist of *ent*-kaur-15-en-17-oic acid (B1), *ent*-kaur-15-en-17,19-dioic acid (B2), *ent*-kaur-15-en-19-al-17-oic acid (B3) and *ent*-kaur-15-en-19-ol-17-oic acid (B4) (Schmelz *et al.* 2011, Block *et al.* 2019, Ding *et al.* 2019). The structural features which differentiate kauralexin A and B compounds is the presence of a double bond at carbon-15 in Kauralexin B compounds. The specific compounds within the kauralexin A and B groups, namely kauralexins A1 to A4 and B1 to B4, are differentiated within the groups by their respective R-groups. The structures of these compounds along with the specific R-groups are depicted below in Figure 1.1.



Kauralexin A

A1: R – CH_3
A2: R – COOH
A3: R – CHO
A4: R – CH_2OH

Kauralexin B

B1: R – CH_3
B2: R – COOH
B3: R – CHO
B4: R – CH_2OH

Figure 1.1: Chemical structures of the two groups of kauralexin secondary metabolites, namely kauralexins A and B. The R-groups which differentiate the individual compounds within the kauralexins groups, namely A1 to A4 and B1 to B4, are shown below the general compound group structures (Poloni and Schirawski 2014, Ding *et al*. 2019).

1.3.3 Role of kauralexins in maize

Kauralexins act as secondary metabolites, forming part of the defence response of maize against biotic and abiotic threats. These compounds have been observed to display antimicrobial and insect antifeedant activity at physiologically relevant concentrations (Harris *et al.* 2005, Schmelz *et al.* 2011, Veenstra *et al.* 2019). Kauralexins have been shown to accumulate in response to fungal pathogen invasion of stem tissue by *Fusarium graminearum, Rhizopus microspores, Colletotrichum ostrophus* and *C. graminicola* (Schmelz *et al.* 2011, Christensen *et al.* 2018). Kauralexin deficient plants were, however, not more susceptible to *C. graminicola* and *F. graminearum* (Christensen *et al.* 2018). In response to insect herbivory, these compounds have been shown to accumulate in response to herbivory by *Ostrinia nubilalis* (Lepidoptera, Crambidae) on maize stem tissue and *Diabrotica balteata* (Coleoptera, Chrysomelidae) on roots (Dafoe *et al.* 2011, Schmelz *et al.* 2011, Vaughan *et al.* 2015). Kauralexins have been shown to play significant roles in maize defences and should be further investigated to determine if these compounds are beneficial against other economically important pests and pathogens.

1.3.4 Anther Ear 2 mutant maize

To investigate the roles of kauralexins in plant defences, a W22 maize inbred line has been produced, which includes a gene-knockout mutation in the *An2* gene (Vaughan *et al.* 2015). The *An2* enzyme is part of the first steps of kauralexins synthesis, providing the necessary precursor compounds to the synthesis pathway (Harris *et al.* 2005, Vaughan *et al.* 2015). The gene-knockout mutation was achieved by using the Activator (Ac) and Dissociation (Ds) system (Ahern *et al.* 2009, Vollbrecht *et al.* 2010). The *AcDs* system is made up of an *Ac* element, an autonomous transposable element, which encodes for a transposase protein, which is able to mobilize the *Ac* and non-autonomous *Ds* elements. The *Ds* element is transposed by the transposase protein, inserting it into the gene of interest and resulting in a gene knock-out (Vollbrecht *et al.* 2010).

The mutant seed was screened to isolate a population with a stable *Ds* insertion in the fourth exon of the *An2* gene, thus producing the mutant. Presence of the mutation was determined using *Ds*-flanking *An2* gene specific primers and by chemotyping the mutant plants to determine the levels of kauralexins present. Plants which contained the *Ds* insertion in both *An2* gene copies (i.e. was homozygous for the mutation) displayed reduced *An2* transcript levels and minimal amounts of kauralexins, this did not result in any morphological defects associated with that of gibberellic acid deficiency (Vaughan *et al.* 2015). Thus far only a few papers have been published in which the *An2* mutant maize line has been used to investigate the role of kauralexins in maize (Table 1.1).

Table 1.1: Published research articles in which the *An2* mutant maize line was used to investigate secondary metabolism in maize.

Title	Main finding	Author / citation
Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance	*An2* mutant was more susceptible to drought. Kauralexin synthesis was positively regulated by drought.	(Vaughan *et al.* 2015)
Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize	*An2* mutant exhibited no production of dolabradienes in the roots whereas W22 maize roots displayed an abundance of dolabradienes.	(Mafu *et al.* 2018)
Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize	*An2* mutant was more susceptible to *Cochliobolus heterostrophus* and *Fusarium verticillioides* compared to W22 maize. No change in susceptibility of the mutant An2 to *Colletotrichum graminicola* or *Fusarium graminearum.*	(Christensen *et al.* 2018)
Kauralexins and zealexins accumulate in sub-tropical maize lines and play a role in seedling resistance to Fusarium verticillioides	*An2* mutant displayed lower levels of kauralexins and zealexins in the roots after *F. verticillioides* inoculation.	(Veenstra *et al.* 2019)

1.4 Lepidopteran Maize Pests in Africa

Insects are incredibly diverse, a recent review estimating the number of different insects being as high as 5.5 million different species globally (Stork 2018). Although there is an abundance of insect species, less than 0.5% of known species are considered pests (Sallam 2013). This small percentage of insect species is said to cause a fifth of the world's annual crop production losses (Sallam 2013). Lepidopteran pests in particular are responsible for the greatest loss of maize crop production globally (Bosque-Pérez 1995). Many lepidopteran pest species have notably reduced maize production in African countries (Table 1.2).

Table 1.2: Major lepidopteran pests of maize in Africa.

Species name	Family	Common name	Region	References
Agrotis ipsilon (Hufnagel)	Noctuidae	Black cutworm	Worldwide	(Odiyo 1975, CABI 2020k)
Busseola fusca (Fuller)	Noctuidae	African maize stalk borer	Sub-Saharan Africa	(Harris and Nwanze 1992, Bosque-Pérez 1995, CABI 2020l)
Chilo partellus (Swinhoe)	Crambidae	Spotted stem borer	Asia, east Africa, central African and southern Africa	(Khadioli *et al.* 2014, CABI 2020m)
Coniesta ignefusalis (Hampson)	Crambidae	Millet stem borer	West Africa and India	(Nwanze and Harris 1992, CABI 2020c)
Eldana saccharina (Walker)	Pyralidae	African sugarcane borer	Africa	(CABI 2020j)
Helicoverpa armigera (Hübner)	Noctuidae	Cotton bollworm	Africa, Europe, Asia, Australia, South America	(Martin *et al.* 2005, CABI 2020g)
Mussidia nigrivenella (Ragonot)	Pyralidae	Maize cob / ear borer	Africa	(Whitney 1970, CABI 2020i)
Sesamia calamistis (Hampson)	Noctuidae	Pink stem borer	Africa	(Van den Berg and Van Wyk 2007, CABI 2020h)
Spodoptera exempta (Walker)	Noctuidae	African armyworm	Africa, South-East Asia, Australasia, Oceania	(Gabriel 1997, CABI 2020a)
Spodoptera exigua (Hübner)	Noctuidae	Lesser armyworm / Beet armyworm	Africa, Europe, Asia, North America, Oceania	(Capinera 1999, Hill 2014, CABI 2020f)
Spodoptera frugiperda (J.E. Smith)	Noctuidae	Fall armyworm	Africa, Americas, Southern Asia, Eastern Asia, Oceania	(Goergen *et al.* 2016, Jacobs *et al.* 2018, CABI 2020e)
Spodoptera littoralis (Boisduval)	Noctuidae	Cotton leafworm	Africa, Europe, Asia	(Salama *et al.* 1971, CABI 2020d)
Thaumatotibia leucotreta (Meyrick)	Tortricidae	False codling moth	Africa and Asia (Israel)	(Catung 1974, Bloem *et al.* 2007, CABI 2020b)

The lepidopteran pests listed above differ in their economic impact in Africa. One of the more notable maize pests, *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae), is considered the most important pest in the highland regions of east and southern Africa, but is of lesser importance in West Africa (Schulthess *et al.* 1997, Ndemah *et al.* 2001, Kfir *et al.* 2002, Sezonlin *et al.* 2006). Of the major lepidopteran maize pests, 30% are *Spodoptera* species, each having a considerable economic impact on the various regions they are most prominent.

1.5 The genus *Spodoptera*

Spodoptera comprises of moth species of the family Noctuidae. This genus is regarded as the most economically devastating group of pests within Noctuidae in agriculture worldwide (Pogue 2002). Members of *Spodoptera* are largely polyphagous and attack crops of great economic importance such as maize, wheat, sugarcane and cotton. In Africa, there are currently four economically important *Spodoptera* species, namely *S. exempta* (African armyworm), *S. exigua* (lesser armyworm), *S. littoralis* (cotton leafworm) and the most recent introduction, *S. frugiperda* (fall armyworm) (Table 1.3).

Table 1.3: Host range and host plant species of *Spodoptera* species present in Africa.

Common name	Species name	Host range	Plant host examples
African armyworm	*Spodoptera exempta* (Walker)	Polyphagous	Field crops (maize, rice, sugarcane, wheat)
Lesser armyworm / Beet armyworm	*Spodoptera exigua* (Hübner)	Polyphagous	Field crops (maize, alfalfa, cotton), vegetable crops (beet, broccoli, cabbage), flower crops (safflower)
Fall armyworm	*Spodoptera frugiperda* (J.E. Smith)	Polyphagous	Field crops (maize, soybean, sorghum, rice, sugarcane), vegetable crops (potato, beans)
Cotton leafworm	*Spodoptera littoralis* (Boisduval)	Polyphagous	Field crops (cotton, maize, onion, tobacco), vegetable crops (beet, cabbage), flower crops (chrysanthemum, amaranthus)

1.5.1 Spodoptera frugiperda

The fall armyworm, *Spodoptera frugiperda,* is a moth which is indigenous to the Americas. The species is found year round in regions of Argentina, in South America, all the way to Southern Florida and Texas, in the United States (Early *et al.* 2018). A quick means of identifying the pest in the field is by observing key morphological characteristics on the adult caterpillar. Two distinctive characteristics are an up-side-down “Y” shape on the head capsule and four spots in the shape of a square on the eight abdominal segment on the caterpillar (Goergen *et al.* 2016) (Figure 1.2).

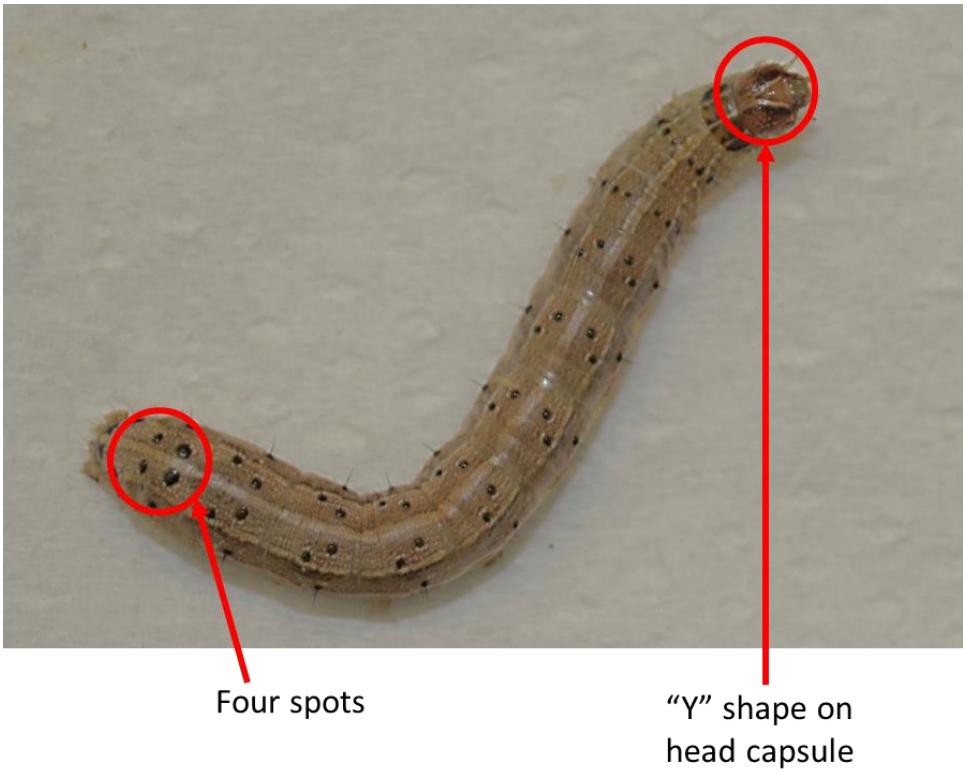


Figure 1.2: Adult fall armyworm caterpillar with key identification characteristics encircled in red, namely the four spots on the last segment of the body and the upside-down "Y" shape on the head capsule. Image: Robert Jansen van Vuuren.

Like the other *Spodoptera* species affecting crops in South Africa, the fall armyworm is also polyphagous in nature, feeding on plant species across multiple different plant families. The species has been recorded from 186 different plant species from 42 plant families that could serve as hosts (Casmuz *et al.* 2010). The vast number of plant species affected by this species includes economically important crops such as maize, rice, sorghum, onion and tomato (Day *et al.* 2017). Although the pest is able to feed on multiple different host plants, host preference is largely determined by the strain of the fall armyworm. Two strains are present within this species, namely the maize strain and the rice strain, exhibiting differences in host preference, chemical resistance, cultivar resistance and mating (Pashley *et al.* 1985, Pashley 1986, Pashley *et al.* 1987, Pashley 1988). The differences between these strains, the composition of the fall armyworm populations in Africa and the impact of these strains on research conducted in Africa is further discussed in Chapter 2 of this study.

1.5.2 Fall armyworm life cycle

The fall armyworm undergoes four life stages throughout its life cycle, an egg stage, typically six larval instars, a pupal stage and the adult moth stage (Vickery 1929). The duration of the life cycle can vary depending on the climatic conditions the insects are exposed to. Under ideal summer conditions, between 21 – 28 °C, the insect will take about 30 days to complete its life cycle, whereas under colder conditions experienced during winter seasons, the life cycle can take as long as 80 – 90 days (Vickery 1929, Capinera 2002).

The eggs are dome shaped and are laid in clusters of approximately 100 to 200 eggs (Luginbill 1928, Capinera 2002). Each of these egg clusters are covered by a layer of scales for protection. Eggs will hatch after about two or three days and the scales covering the eggs will serve as the first food source for the newly hatched larvae. Newly hatched larvae (neonates), will feed on available foliage, grow and transition through the six typical instar stages. The duration of these instar stages is summarised in Table 1.4.

Table 1.4: Fall armyworm larval instar stage durations in hours.

Instar stage	Minimum (hours)	Maximum (hours)	Average (hours)
1	45	57	47.9
2	20	50	34.6
3	20	50	31.2
4	13	44	32.3
5	36	75.5	51.1
6	49	119	93.2

**Data represented in this table was acquired from (Luginbill 1928).*

Larvae in the first or second instar are whitish in colour with no spines along the body and a black head capsule. As the larvae mature to late third and beginning fourth instars, the head capsules change to a dark orange colour, black spines are present along the body and the body changes to a brown colour with white lateral lines (Capinera 2002). Between instars four and six, the head capsule is a red brown colour with body's that are brown with dorsal white stripes. In the later instar stages dark spots become more apparent on the bodies of the larvae (Capinera 2002). The identifiable markers, namely the "Y" shape on the head capsule and four spots on the last segment of the larvae are present as early as the second instar (Figure 1.2). The larval stage of the pest is the most destructive stage as they feed on both vegetative and reproductive structures of the plant hosts (Day *et al.* 2017). In the case of maize, overall yield is negatively influenced as foliage is consumed, reducing photosynthetic capacity of the plants, and kernels are consumed, directly reducing marketable yield (Day *et al.* 2017). Examples of the damage they cause on maize plants can be seen below in Figure 1.3.

Figure 1.3: Fall armyworm feeding damage on different part of the maize crop, namely the maize stem (A), foliage (B) and maize cob (C). Images: A – (Goergen *et al.* 2016), B – (Ufumeli 2019), C – (www.bayer.com).

The sixth instar, once ready to pupate will fall to the ground and burrow about 2,5 to 7,6 cm deep into the soil (Vickery 1929). The depth burrowed is likely dependant on soil moisture and soil type (Shi *et al.* 2021). The larvae then pupates in the soil and the pupal stage will continue for about seven to thirty-seven days depending on soil temperature (Vickery 1929). The adult moths emerge during the first few hours of the evening, the newly moulted adult displays a reluctance to fly and does not mate the first evening, but only feeds on pollen and floral nectar (Sparks 1979, He *et al.* 2021). The forewings of male moths are generally shaded grey and brown, accompanied by white triangular markings at the tips and centre of the wings. The female moths do not have distinctive markings but display the same brown and grey colours (Capinera 2002). The moths are nocturnal, beginning their movement towards suitable hosts during the early evening, these hosts serve as sites for feeding, mating and oviposition. Female moths attract male moths for mating by releasing sex pheromones while being perched on a suitable plant host. Males have been shown to respond to these signals as far as 9 – 12 m away, depending on wind speed and temperature (Sparks 1979). Competition between males is high as females only mate once in an evening. Mating activities of these moths have been shown to peak during midnight, with oviposition following close after mating. In areas where fall armyworm population density is low, the females will lay their eggs on the underside of leaves, whereas in areas where the fall armyworm population density is high the females will lay their eggs on any available surface, be this on different plants organs or even man-made structures such as window panes (Sparks 1979). The average life span of an adult moth varies between 10 and 21 days (Capinera 2002).

1.5.3 The fall armyworm in Africa

The first fall armyworm case reported in Africa in 2016 came from West and Central Africa, namely Nigeria, Benin and São Tomé. The morphological identification of the pest was confirmed by sequencing the *Cytochrome oxidase I* (*COI*) gene, which then also confirmed the presence of both strains in West and Central Africa (Goergen *et al.* 2016). Fall armyworm

was then reported in Ghana, it is speculated that this introduction came via West Africa. Both strains were confirmed to be present in Ghana by sequencing *COI*, with both strains being found on maize and no data at this point suggesting a wider host range in Africa (Cock *et al.* 2017). In 2017, the fall armyworm was also confirmed to be present in Togo, which is situated between Ghana and Benin. Most of the devastation caused by the fall armyworm in Africa has been on crop plants associated with the maize strain, namely maize and sorghum, and rice strain preferred hosts (i.e. rice and Bermuda grass) having been mostly unaffected (Nagoshi 2019).

The fall armyworm spread rapidly in Africa. Presence of the pest was confirmed in southern Africa and specifically South Africa by 2017 (Jacobs *et al.* 2018). Surveys of local farmers in Zambia and Ghana have reported an estimated mean maize yield loss of 45% and 40% respectively due to FAW (Day *et al.* 2017). Similarly, surveys were conducted on smallholder maize farmers from Ethiopia and Kenya, estimating yield losses to be averages of 32% and 47.3%, respectively (Kumela *et al.* 2019). It is important to note however that these surveys were based on the perceptions of the farmers affected and were not rigorous field evaluations. When field evaluations were conducted by Baudron *et al.* in 2019 on smallholder maize fields in eastern Zimbabwe, fields which exhibited fall armyworm infestation and damage only resulted in an average yield loss of 11.57% (Baudron *et al.* 2019). In Ethiopia, the average yield loss of maize was also determined to be 11.5% (Kassie *et al.* 2020), this yield loss was after various control measures were implemented. The reason for the discrepancies between the farmer perceptions and field evaluation results concerning yield loss due to the fall armyworm is unclear. The yield losses observed by farmers recorded through surveys may have included losses due to various environmental factors but this cannot be confirmed.

1.6 Current Fall Armyworm Control Strategies

1.6.1 Resistant transgenic maize varieties

Transgenic maize which produces *Bacillus thurigiensis* (*Bt*) toxins is being utilised all over the world for insect pest control. These *Bt* toxins have been effective against various insect pests such as armyworms, stem borers and maize root- and earworms. The commercialization of *Bt* genetically modified crops took place from 1996 to 2002, these GM crops were planted globally reaching a cumulative total of over 235 million hectares (James 2011). In 2010 alone, more than 58 million hectares of transgenic *Bt* maize and *Bt* cotton were planted worldwide (James 2011).

There have been multiple accounts however of insect resistance developing against *Bt* crops, challenging the efficacy and driving the further development of these transgenic crops (Tabashnik *et al.* 2008, Kruger *et al.* 2009, Tabashnik *et al.* 2013). The fall armyworm has been a major target of *Bt* maize in both North and South America, but resistance to many of these *Bt* maize transgenic lines has been noted due to their extensive use (Frizzas *et al.* 2014, Huang *et al.* 2014). The Cyr1F *Bt* maize is an example of one of these *Bt* transgenic lines. Fall armyworm populations containing multiple resistance alleles to Cry1F *Bt* maize have been reported in south-eastern regions of the US and in South America, particularly Puerto Rico (Huang *et al.* 2014). In South Africa, a similar problem is being experienced. Fall armyworm populations in South Africa have displayed moderate survival rates on *Bt* maize which contains single-toxin events, such as Cry1Ab. However, high larval mortality rates have been observed on *Bt* maize which contains pyramid-toxin events, such as Cry1A.105 + Cry2Ab2 (Botha *et al.* 2019). The presence of resistance alleles in the African fall armyworm population is speculated (Botha *et al.* 2019). The high mortality of these fall armyworm populations on pyramid-toxin event *Bt* maize strains is a promising control. The use of multiple *Bt* toxins within maize

transgenic lines is beneficial as the presence of multiple different *Bt* toxins within a transgenic maize line means that the pest would require multiple resistance alleles to overcome the plants defences (Carrière *et al.* 2016).

1.6.2 Synthetic pesticides

Early instar fall armyworm larvae tend hide in the maize funnel during the day, coming out more in the evening to feed on the leaves. This means that pesticide sprays need to be applied at either dawn or dusk to be effective (Day *et al.* 2017). Older, more veracious larvae tend to remain close to the maize stem, staying in the maize funnel, rendering pesticide sprays less effective. In Africa, mostly cheaper pesticides are used. These pesticides include pesticides mainly consisting of organophosphates and pyrethroids-pyrethrins, both of which FAW larvae in the Americas have developed resistance against (Day *et al.* 2017).

When surveying smallholder maize farmers across five African countries that have the pest, it was noted that the most popular control measure was synthetic pesticides. The insecticide usage and frequency of usage dramatically increased with identification of this pest within fields, pesticide usage could be predicted to increase by 13% and frequency of usage by 38% (Kassie *et al.* 2020). The reason for this could be linked to the lack of extension work done in many of the relevant communities and the large government subsidies which allow for greater access to these pesticides (Tambo *et al.* 2020). The extensive use of these various synthetic pesticides is potentially hazardous to human health, especially as it has been noted that many farmers would handle these chemicals without the proper protective attire (Tambo *et al.* 2020).

1.6.3 Biological control

Biological control of fall armyworm larvae in the form of natural predators and parasitoids have been noted to occur in the tropical and subtropical regions of the Americas (Ashley 1979). Three typical strategies can be employed to combat an insect pest using its natural enemies (Van Driesche and Bellows 1996). The first being classical biological control in which natural enemies are mass produced and released to take up permanent residence. The second being augmentative biological control in which the natural enemies are mass produced to be released regularly for temporary control. The last strategy is conservation biological control, which involves conserving natural enemies that are already present and enhancing their efficiency by using selective insecticides.

In the Americas, a total of fifty-three parasitoid species were identified in a study conducted by Ashley in 1979. These species made up 43 genera and 10 families, thus showing the large number of parasitoids that are able to parasitise the FAW (Ashley 1979). The two parasitoids most frequently recovered in this survey were *Apanteles marginiventris* (Cresson) and *Chelonus texanus* (Cresson). In South America, maize plants it was noted that maize plants could be treated with a sugar solution to attract parasitoids and natural predators of the fall armyworm. This technique resulted in reduced levels of fall armyworm infestation on maize (18% lower than untreated plants) and reduced leaf area damage by an average of 35% (Canas and O'Neil 1998). This study demonstrated the impact natural enemies can have in pest control.

Naturally occurring parasitoids in Africa have been noted to parasitise the FAW. In a study conducted by Sisay *et al.* (2018), the occurrence and identity of these parasitoids were determined in three African countries (Ethiopia, Kenya and Tanzania) with reported FAW presence. Farms in these countries, were surveyed in regions of known FAW presence by collecting FAW larvae, eggs and adult moths. These specimens were allowed to develop in a

laboratory under sterile conditions, recording the number of parasitoids, which emerged from the FAW specimens. In Ethiopia, the most prominent larval parasitoid was *Cotesia icipe* (Fernandez-Triana & Fiobe) (Hymenoptera: Braconidae), which occurred at three different farms with a 33.8% to 45.3% parasitism level. Two other parasitoids emerged from the larvae, namely *Palexorista zonata* (Curran) and *Coccygidium luteum* (Brullé). Two parasitoid wasps reported in Kenya, namely *Charops ater* Szépligeti and *P. zonata,* both caused roughly 12% parasitism of the collected samples in the relevant regions (Sisay *et al.* 2018). In Tanzania, *C. ater* was the most notable larval parasitoid, causing 10% parasitism. The discovery of these new associations of known, natural parasitoids with the alien pest, the fall armyworm, holds promise for the development of integrated pest management strategies utilising these parasitoids (Sisay *et al.* 2018).

It is interesting to note however that the use of *Bt* toxin maize may impede the effectivity of parasitoid related biological control of this pest. In a study conducted by Desneux *et al*. (2010), it was discovered that the lepidopteran parasitoid *Cotesia marginiventris* exhibited significantly weaker responses to *S. frugiperda* frass from larvae that fed on *Bt* maize as compared to larvae fed on conventional maize (Desneux *et al.* 2010). This means that the parasitoid exhibited a weaker attraction to the larvae feeding on *Bt* maize, and thus may lower the possibility of the parasitoid detecting the fall armyworm larvae. This result was thought to be due to *Bt* toxins having a detrimental effect on the larval gut microbiome (Desneux *et al.* 2010). Although frass from larvae which fed on either *Bt* or conventional maize both still resulted in attraction of the parasitoid, it is still important to note as these unintended outcomes may impact the efficiency of natural biological control.

1.6.4 Cultural control strategies

Cultural or physical control strategies are popular means of controlling the fall armyworm in many African countries. Some of these strategies include the physical removal of caterpillars and egg masses from crop plants, pouring ash and/or sand into the whorls of maize plants, frequent weeding, intercropping with non-host plants and the use of push-pull intercropping (Midega *et al.* 2018, Tambo *et al.* 2020). The physical removal of caterpillars and egg masses has been used by many small holder farmers where the small size of the farms allow for this more labour-intensive method. Frequent weeding of potential alternative hosts, such as pasture grasses, and intercropping with non-host crops such as cassava have also been used (Tambo *et al.* 2020). Climate-adapted push-pull intercropping has shown to be a highly effective control strategy. This strategy involves intercropping maize with *Desmodium intortum* and planting *Brachiaria* cv Mulato II as a border crop. The desmodium plants emit semiochemicals which repel the lepidopteran pests, while *Brachiaria*, the border crop, attracts the pests. In Kenya, Uganda and Tanzania, this method reduced the number of fall armyworm larvae observed on maize crops by an average of about 82% and reduced crop damage by an average of about 86% as compared to monocrop fields (Midega *et al.* 2018). Push-pull intercropping is a valuable control strategy and should be considered when developing an integrated pest management plan for smallholder farmers in Africa (Khan *et al.* 2018).

1.7 Bioassays to determine the effect of phytochemicals

Insect feeding assays have been integral in determining the efficiency of certain plant defence compounds in the hopes of developing new insecticides or more resistant transgenic crops. Parameters such as larval mass gain, developmental time and total plant material consumed are simple parameters in these feeding assays which can be directly linked to crop protection (Knolhoff and Heckel 2014). Non-choice feeding assays involve introducing insect specimens to plant material / artificial media, which contains or has been treated with the compounds of interest while providing no other food alternative. This type of feeding assay aims to determine whether the compounds of interest display insect antifeedant activity, either reducing or inhibiting insect feeding (Munakata 1975, Isman 2002), or are toxic, influencing the mortality of the insects directly. Choice feeding assays involve introducing insect specimens to plant material / artificial media which contains or has been treated with the compounds of interest and plant material / artificial media which does not. This type of feeding assay aims to determine whether the compounds of interest influence the feeding preference of the insect by acting as a deterrent (Szczepanik *et al.* 2005).

An example of a non-choice feeding assay is present in a study by Pavela in 2010. The study investigated the antifeedant activity of plant extracts from 75 different plant species of the Eurasian region against *S. littorallis.* In these assays, plant extractions of varying concentrations were applied to leaf discs and larvae were allowed to feed on single leaf discs. These experiments highlighted four extracts, which were the most effective in deterring larval feeding. The polyphagous nature of the pest was listed as a potential reason for why so few of the extracts were effective (Pavela 2010).

In another study, chemically altered terpenoid lactones were tested for feeding deterrent activity against the Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Coleoptera, Chrysomelidae), by employing non-choice assays and choice assays (Szczepanik *et al.* 2005). The feeding assays involved treating leaf discs with these different terpenoid lactones and allowing different developmental stages of the insect to feed on one leaf disc (non-choice assay) or to choose between a treated leaf disc or a control (choice assay). The non-choice assays determined that some of these terpenoid lactones exhibits antifeedant activity, resulting in lower leaf tissue consumption. The choice assay data revealed that many of those same terpenoid lactones also exhibit strong deterrent properties which influence the beetles' preference (Szczepanik *et al.* 2005). Feeding assays are a relatively simple way to test whether certain phytochemicals have antifeedant, deterrent or toxic properties which could be further investigated for the purpose of crop protection.

1.8 References

Aharoni, A., M. A. Jongsma, and H. J. Bouwmeester. 2005. Volatile science? Metabolic engineering of terpenoids in plants. Trends in Plant Science 10: 594-602.

Ahern, K. R., P. Deewatthanawong, J. Schares, M. Muszynski, R. Weeks, E. Vollbrecht, J. Duvick, V. P. Brendel, and T. P. Brutnell. 2009. Regional mutagenesis using dissociation in maize. Methods 49: 248-254.

Ashley, T. R. 1979. Classification and distribution of fall armyworm parasites. Florida Entomologist: 62:114-123.

Baudron, F., M. A. Zaman-Allah, I. Chaipa, N. Chari, and P. Chinwada. 2019. Understanding the factors influencing fall armyworm (*Spodoptera frugiperda* JE Smith) damage in African smallholder maize fields and quantifying its impact on yield. A case study in Eastern Zimbabwe. Crop Protection 120: 141-150.

Bebber, D. P., T. Holmes, D. Smith, and S. J. Gurr. 2014. Economic and physical determinants of the global distributions of crop pests and pathogens. New Phytologist 202: 901-910.

Bender, R. R., J. W. Haegele, M. L. Ruffo, and F. E. Below. 2013. Nutrient uptake, partitioning, and remobilization in modern, transgenic insect-protected maize hybrids. Agronomy Journal 105: 161-170.

Bennetzen, J., E. Buckler, V. Chandler, J. Doebley, J. Dorweiler, B. Gaut, M. Freeling, S. Hake, E. Kellogg, and R. S. Poethig. 2001. Genetic evidence and the origin of maize. Latin American Antiquity 12: 84-86.

Block, A. K., M. M. Vaughan, E. A. Schmelz, and S. A. Christensen. 2019. Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). Planta 249: 21-30.

Bloem, S., J. Carpenter, and H. Hofmeyr. 2007. Area-wide control tactics for the false codling moth *Thaumatotibia leucotreta* in South Africa: a potential invasive species, pp. 351-359, Area-Wide Control of Insect Pests. Springer.

Bosque-Pérez, N. A. 1995. Major insects pests of maize in Africa: biology and control. IITA research guide, No. 30. Ibadan, Nigeria: IITA, (p. 30). https://hdl.handle.net/10568/97254

Botha, A. S., A. Erasmus, H. du Plessis, and J. Van den Berg. 2019. Efficacy of Bt maize for control of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in South Africa. Journal of Economic Entomology 112: 1260-1266.

CABI. 2020a. *Chilo partellus* (spotted stem borer), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020b. *Spodoptera exigua* (beet armyworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020c. *Spodoptera exempta* (black armyworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020d. *Helicoverpa armigera* (cotton bollworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020e. *Coniesta ignefusalis*, Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020f. *Spodoptera frugiperda* (fall armyworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020g. *Eldana saccharina* (African sugarcane borer), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020h. *Spodoptera littoralis* (cotton leafworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020i. *Sesamia calamistis* (African pink stem borer), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020j. *Mussidia nigrivenella* (cob borer), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020k. *Thaumatotibia leucotreta* (false codling moth), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020l. *Agrotis ipsilon* (black cutworm), Invasive Species Compendium. CABI International, Wallingford, UK.

CABI. 2020m. *Busseola fusca* (African maize stalk borer), Invasive Species Compendium. CABI International, Wallingford, UK.

Canas, L. A., and R. J. O'Neil. 1998. Applications of sugar solutions to maize, and the impact of natural enemies on fall armyworm. International Journal of Pest Management 44: 59-64.

Capinera, J. L. 1999. Beet Armyworm, *Spodoptera exigua* (Hübner)(Insecta: Lepidoptera: Noctuidae). EDIS.

Capinera, J. L. 2002. Fall armyworm, *Spodoptera frugiperda* (JE Smith)(Insecta: Lepidoptera: Noctuidae). EDIS 7.

Carena, M., G. Bergman, N. Riveland, E. Eriksmoen, and M. Halvorson. 2009. Breeding maize for higher yield and quality under drought stress. Maydica 54: 287.

Carrière, Y., J. A. Fabrick, and B. E. Tabashnik. 2016. Can pyramids and seed mixtures delay resistance to Bt crops? Trends in Biotechnology 34: 291-302.

Casmuz, A., M. L. Juárez, M. G. Socías, M. G. Murúa, S. Prieto, S. Medina, E. Willink, and G. Gastaminza. 2010. Revisión de los hospederos del gusano cogollero del maíz,

Spodoptera frugiperda (Lepidoptera: Noctuidae). Revista de la Sociedad Entomológica Argentina 69: 209-231.

Catung, H. 1974. Population studies of the false codling moth, *Cryptophlebia lellcolrela* Meyr., on citrus in the Transvaal. Phytophylactica 6: 31-38.

Christensen, S. A., J. Sims, M. M. Vaughan, C. Hunter, A. Block, D. Willett, H. T. Alborn, A. Huffaker, and E. A. Schmelz. 2018. Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize. Journal of Experimental Botany 69: 1693-1705.

Cock, M. J., P. K. Beseh, A. G. Buddie, G. Cafá, and J. Crozier. 2017. Molecular methods to detect *Spodoptera frugiperda* in Ghana, and implications for monitoring the spread of invasive species in developing countries. Scientific Reports 7: 1-10. DOI: 10.1038/s41598-017-04238-y

Crow, J. F. 1998. 90 years ago: the beginning of hybrid maize. Genetics 148: 923-928.

Dafoe, N. J., A. Huffaker, M. M. Vaughan, A. J. Duehl, P. E. Teal, and E. A. Schmelz. 2011. Rapidly induced chemical defenses in maize stems and their effects on short-term growth of *Ostrinia nubilalis*. Journal of Chemical Ecology 37: 984-991.

Day, R., P. Abrahams, M. Bateman, T. Beale, V. Clottey, M. Cock, Y. Colmenarez, N. Corniani, R. Early, and J. Godwin. 2017. Fall armyworm: impacts and implications for Africa. Outlooks on Pest Management 28: 196-201.

Desneux, N., R. Ramírez-Romero, A. H. Bokonon-Ganta, and J. S. Bernal. 2010. Attraction of the parasitoid *Cotesia marginiventris* to host (*Spodoptera frugiperda*) frass is affected by transgenic maize. Ecotoxicology 19: 1183-1192.

Ding, Y., K. M. Murphy, E. Poretsky, S. Mafu, B. Yang, S. N. Char, S. A. Christensen, E. Saldivar, M. Wu, Q. Wang, L. Ji, R. J. Schmitz, K. A. Kremling, E. S. Buckler, Z. Shen, S. P. Briggs, J. Bohlmann, A. Sher, G. Castro-Falcon, C. C. Hughes, A. Huffaker, P. Zerbe, and E. A. Schmelz. 2019. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. Nature Plants 5: 1043-1056.

Early, R., P. González-Moreno, S. T. Murphy, and R. Day. 2018. Forecasting the global extent of invasion of the cereal pest *Spodoptera frugiperda*, the fall armyworm. NeoBiota 40: 25-50.

Eubanks, M. W. 2001. The mysterious origin of maize. Economic Botany: 55:492-514. https://www.jstor.org/stable/4256485

Fakorede, M., B. Badu-Apraku, A. Kamara, A. Menkir, and S. Ajala. 2003. Maize revolution in West and Central Africa: an overview. Maize revolution in West and Central Africa: 3-15.

FAO. 2019. Production Quantity, FAOSTAT Statistical Database. Food and Agriculture Organization of the United Nations.

Frizzas, M. R., S. Silveira Neto, C. M. d. Oliveira, and C. Omoto. 2014. Genetically modified corn on fall armyworm and earwig populations under field conditions. Ciencia Rural 44: 203-209.

Gabriel, B. P. 1997. Insects and mites injurious to Philippine crop plants, vol. 171, National Crop Protection Center.

Goergen, G., P. L. Kumar, S. B. Sankung, A. Togola, and M. Tamo. 2016. First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae), a new alien invasive pest in West and Central Africa. PloS One 11: e0165632.

Harris, K. M., and K. Nwanze. 1992. *Busseola fusca* (Fuller), the African maize stalk borer: a handbook of information, International Crops Research Institute for the Semi-Arid Tropics.

Harris, L. J., A. Saparno, A. Johnston, S. Prisic, M. Xu, S. Allard, A. Kathiresan, T. Ouellet, and R. J. Peters. 2005. The maize *An2* gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. Plant Molecular Biology 59: 881-894.

He, L.-m., S. Jiang, Y.-c. Chen, K. A. Wyckhuys, S.-s. Ge, H. Wei, X.-w. Gao, and K.-m. Wu. 2021. Adult nutrition affects reproduction and flight performance of the invasive fall armyworm, Spodoptera frugiperda in China. Journal of Integrative Agriculture 20: 715-726.

Hill, L. 2014. Lesser armyworm,'*Spodoptera exigua*'(Hubner)(Lepidoptera: Noctuidae), a vagrant moth in Tasmania. Plant Protection Quarterly 29: 131.

Howe, G. A., and G. Jander. 2008. Plant immunity to insect herbivores. Annual Review of Plant Biology. 59: 41-66.

Huang, F., J. A. Qureshi, R. L. Meagher Jr, D. D. Reisig, G. P. Head, D. A. Andow, X. Ni, D. Kerns, G. D. Buntin, and Y. Niu. 2014. Cry1F resistance in fall armyworm *Spodoptera frugiperda*: single gene versus pyramided Bt maize. PloS One 9: e112958.

Isman, M. 2002. Insect antifeedants. Pesticide Outlook 13: 152-157.

Jacobs, A., A. V. Vuuren, and I. H. Rong. 2018. Characterisation of the fall armyworm (*Spodoptera frugiperda* J. E. Smith) (Lepidoptera: Noctuidae) from South Africa. African Entomology 26: 45-49.

James, C. 2003. Global review of commercialized transgenic crops: 2002 feature: Bt maize, vol. 29, ISAAA Ithaca, NY.

James, C. 2011. Global status of commercialized biotech/GM crops, 2011. vol. 44, ISAAA Ithaca, NY.

Kassie, M., T. Wossen, H. De Groote, T. Tefera, S. Sevgan, and S. Balew. 2020. Economic impacts of fall armyworm and its management strategies: evidence from southern Ethiopia. European Review of Agricultural Economics 47: 1473-1501.

Kfir, R., W. Overholt, Z. Khan, and A. Polaszek. 2002. Biology and management of economically important lepidopteran cereal stem borers in Africa. Annual Review of Entomology 47: 701-731.

Khadioli, N., Z. Tonnang, E. Muchugu, G. Ong'amo, T. Achia, I. Kipchirchir, J. Kroschel, and B. Le Ru. 2014. Effect of temperature on the phenology of *Chilo partellus* (Swinhoe)(Lepidoptera, Crambidae); simulation and visualization of the potential future distribution of *C. partellus* in Africa under warmer temperatures through the development of life-table parameters. Bulletin of Entomological Research 104: 809-822.

Khan, Z. R., J. O. Pittchar, C. A. Midega, and J. A. Pickett. 2018. Push-pull farming system controls fall armyworm: lessons from Africa. Outlooks on Pest Management 29: 220-224.

Knolhoff, L. M., and D. G. Heckel. 2014. Behavioral assays for studies of host plant choice and adaptation in herbivorous insects. Annual Review of Entomology 59: 263-278.

Kruger, M., J. B. J. Van Rensburg, and J. Van den Berg. 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: 684-689.

Kumela, T., J. Simiyu, B. Sisay, P. Likhayo, E. Mendesil, L. Gohole, and T. Tefera. 2019. Farmers' knowledge, perceptions, and management practices of the new invasive pest, fall armyworm (*Spodoptera frugiperda*) in Ethiopia and Kenya. International Journal of Pest Management 65: 1-9. https://doi.org/10.1080/09670874.2017.1423129

Lee, M. 1994. Inbred lines of maize and their molecular markers, pp. 423-432, The Maize Handbook. Springer, New York, NY.

Luginbill, P. 1928. The fall army worm, vol. 34, US Department of Agriculture.

Mahmoud, S. S., and R. B. Croteau. 2002. Strategies for transgenic manipulation of monoterpene biosynthesis in plants. Trends in Plant Science 7: 366-373.

Martin, T., G. O. Ochou, A. Djihinto, D. Traore, M. Togola, J. M. Vassal, M. Vaissayre, and D. Fournier. 2005. Controlling an insecticide-resistant bollworm in West Africa. Agriculture, Ecosystems & Environment 107: 409-411.

Meyer, J., S. L. Murray, and D. K. Berger. 2016. Signals that stop the rot: regulation of secondary metabolite defences in cereals. Physiological and Molecular Plant Pathology 94: 156-166.

Midega, C. A., J. O. Pittchar, J. A. Pickett, G. W. Hailu, and Z. R. Khan. 2018. A climate-adapted push-pull system effectively controls fall armyworm, *Spodoptera frugiperda* (JE Smith), in maize in East Africa. Crop protection 105: 10-15.

Munakata, K. 1975. Insect antifeeding substances in plant leaves, pp. 57-66, Pesticide Chemistry–3. Elsevier.

Murphy, K. M., L.-T. Ma, Y. Ding, E. A. Schmelz, and P. Zerbe. 2018. Functional characterization of two class II diterpene synthases indicates additional specialized diterpenoid pathways in maize (*Zea mays*). Frontiers in Plant Science 9: 1542. https://doi.org/10.3389/fpls.2018.01542

Nagoshi, R. N. 2019. Evidence that a major subpopulation of fall armyworm found in the Western Hemisphere is rare or absent in Africa, which may limit the range of crops at risk of infestation. PloS One 14. https://doi.org/10.1371/journal.pone.0208966

Ndemah, R., F. Schulthess, S. Korie, C. Borgemeister, and K. Cardwell. 2001. Distribution, relative importance and effect of lepidopterous borers on maize yields in the forest zone and mid-altitude of Cameroon. Journal of Economic Entomology 94: 1434-1444.

Nwanze, K. F., and K. Harris. 1992. Insect pests of pearl millet in West Africa. Review of Agriculture Entomology 80: 1132-1155.

Odiyo, P. O. 1975. Seasonal distribution and migrations of *Agrotis ipsilon* (Hufnagel)(Lepidoptera, Noctuidae), Centre for Overseas Pest Research (COPR).

Pashley, D. P. 1986. Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): a sibling species complex? Annals of the Entomological Society of America 79: 898-904.

Pashley, D. P. 1988. Quantitative genetics, development, and physiological adaptation in host strains of fall armyworm. Evolution 42: 93-102.

Pashley, D. P., S. J. Johnson, and A. N. Sparks. 1985. Genetic population structure of migratory moths: the fall armyworm (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 78: 756-762.

Pashley, D. P., S. S. Quisenberry, and T. Jamjanya. 1987. Impact of fall armyworm (Lepidoptera: Noctuidae) host strains on the evaluation of bermuda grass resistance. Journal of Economic Entomology 80: 1127-1130.

Pavela, R. 2010. Antifeedant activity of plant extracts on *Leptinotarsa decemlineata* Say. and *Spodoptera littoralis* Bois. larvae. Industrial Crops and Products 32: 213-219.

Pogue, M. G. 2002. A world revision of the genus *Spodoptera* Guenée:(Lepidoptera: Noctuidae). Memoirs of the American Entomological Society: 1-202.

Poloni, A., and J. Schirawski. 2014. Red card for pathogens: phytoalexins in sorghum and maize. Molecules 19: 9114-9133.

Pratt, R. C., and S. G. Gordon. 2006. Breeding for resistance to maize foliar pathogens, vol. 27, Wiley.

Ranum, P., J. P. Peña-Rosas, and M. N. Garcia-Casal. 2014. Global maize production, utilization, and consumption. Annals of the New York Academy of Sciences 1312: 105-112.

Rodríguez-Concepción, M., and A. Boronat. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology 130: 1079-1089.

Salama, H., N. Z. Dimetry, and S. Salem. 1971. On the host preference and biology of the cotton leaf worm *Spodoptera littoralis* Bois. Zeitschrift für Angewandte Entomologie 67: 261-266.

Sallam, M. N. 2013. Insect damage: damage on post-harvest, Food and Agriculture Organization of the United Nations.

Schmelz, E. A., F. Kaplan, A. Huffaker, N. J. Dafoe, M. M. Vaughan, X. Ni, J. R. Rocca, H. T. Alborn, and P. E. Teal. 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. Proceedings of the National Academy of Sciences 108: 5455-5460.

Schulthess, F., N. Bosque-Perez, A. Chabi-Olaye, S. Gounou, R. Ndemah, and G. Goergen. 1997. Exchange of natural enemies of Lepidofteran cereal stemborers between african regions. International Journal of Tropical Insect Science 17: 97-108.

Sezonlin, M., S. Dupas, B. Le Rü, P. Le Gall, P. Moyal, P. A. Calatayud, I. Giffard, N. Faure, and J. F. Silvain. 2006. Phylogeography and population genetics of the maize stalk borer *Busseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa. Molecular Ecology 15: 407-420.

Shi, Y., L.-Y. Li, S. Shahid, G. Smagghe, and T.-X. Liu. 2021. Effect of soil moisture on pupation behavior and inhabitation of Spodoptera frugiperda (Lepidoptera: Noctuidae). Applied Entomology and Zoology 56: 69-74.

Sisay, B., J. Simiyu, P. Malusi, P. Likhayo, E. Mendesil, N. Elibariki, M. Wakgari, G. Ayalew, and T. Tefera. 2018. First report of the fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae), natural enemies from Africa. Journal of Applied Entomology 142: 800-804.

Sluyter, A., and G. Dominguez. 2006. Early maize (*Zea mays* L.) cultivation in Mexico: Dating sedimentary pollen records and its implications. Proceedings of the National Academy of Sciences 103: 1147-1151.

Sparks, A. N. 1979. A review of the biology of the fall armyworm. Florida Entomologist: 82-87.

Stork, N. E. 2018. How many species of insects and other terrestrial arthropods are there on Earth? Annual Review of Entomology 63: 31-45.

Șuteu, D., I. Băcilă, V. Haș, I. Haș, and M. Miclăuș. 2013. Romanian Maize (*Zea mays*) inbred lines as a source of genetic diversity in SE Europe, and their potential in future breeding efforts. PloS One 8: e85501.

Szczepanik, M., I. Dams, and C. Wawrzeńczyk. 2005. Feeding deterrent activity of terpenoid lactones with the *p*-menthane system against the Colorado potato beetle (Coleoptera: Chrysomelidae). Environmental Entomology 34: 1433-1440.

Tabashnik, B. E., T. Brévault, and Y. Carrière. 2013. Insect resistance to Bt crops: lessons from the first billion acres. Nature Biotechnology 31: 510-521.

Tabashnik, B. E., A. J. Gassmann, D. W. Crowder, and Y. Carrière. 2008. Insect resistance to Bt crops: evidence versus theory. Nature Biotechnology 26: 199-202.

Tambo, J. A., M. K. Kansiime, I. Mugambi, I. Rwomushana, M. Kenis, R. K. Day, and J. Lamontagne-Godwin. 2020. Understanding smallholders' responses to fall armyworm (*Spodoptera frugiperda*) invasion: Evidence from five African countries. Science of the Total Environment 740: 140015.

Ufumeli, A. 2019. African countries should turn to lower risk solutions to fight fall armyworm.

Van den Berg, J., and A. Van Wyk. 2007. The effect of Bt maize on *Sesamia calamistis* in South Africa. Entomologia Experimentalis et Applicata 122: 45-51.

Van Driesche, R. G., and T. S. Bellows. 1996. Biology of arthropod parasitoids and predators, pp. 309-336, Biological Control. Springer, Boston, MA.

Vaughan, M. M., S. Christensen, E. A. Schmelz, A. Huffaker, H. J. Mcauslane, H. T. Alborn, M. Romero, L. H. Allen, and P. E. A. Teal. 2015. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. Plant, Cell & Environment 38: 2195-2207.

Veenstra, A., N. Moola, S. Wighard, J. Korsman, S. A. Christensen, M. S. Rafudeen, and S. L. Murray. 2019. Kauralexins and zealexins accumulate in sub-tropical maize lines and play a role in seedling resistance to *Fusarium verticillioides*. European Journal of Plant Pathology 153: 223-237.

Vickery, R. A. Year. Published. Studies on the fall army worm in the gulf coast district of Texas. *In*, 1929.

Vollbrecht, E., J. Duvick, J. P. Schares, K. R. Ahern, P. Deewatthanawong, L. Xu, L. J. Conrad, K. Kikuchi, T. A. Kubinec, and B. D. Hall. 2010. Genome-wide distribution of transposed dissociation elements in maize. The Plant Cell 22: 1667-1685.

Whitney, W. 1970. Observations on maize insects at the International Institute of Tropical Agriculture (IITA) Ibadan. Bulletin of the Entomological Society of Nigeria 2: 146-155.

Yamane, H. 2013. Biosynthesis of phytoalexins and regulatory mechanisms of it in rice. Bioscience, Biotechnology, and Biochemistry 77: 1141-1148.

Zampieri, M., A. Ceglar, F. Dentener, A. Dosio, G. Naumann, M. Van Den Berg, and A. Toreti. 2019. When will current climate extremes affecting maize production become the norm? Earth's Future 7: 113-122.

Chapter 2: The importance of determining strain identity in a laboratory reared fall armyworm culture

2.1 Abstract

The fall armyworm is an insect pest which is comprised of two strains, namely the corn and rice strain. These strains, typically defined by the *COI* mitochondrial and *Tpi* sex-linked markers, have been shown to be present in Africa. Further investigation has revealed that the majority of the population in Africa is comprised of interstrain hybrids, the impact that this may have on host preference and the efficacy of specific control strategies is unknown. The genomes of the corn and rice strain have been shown to be significantly different in detoxification and gustatory genes, which may explain their differences in host preferences and insecticidal resistance. With the high prevalence of interstrain hybridisation, these nuclear genes may be exchanged, resulting in experiments which yield irreplicable results and the development of ineffective control strategies. In this study, fall armyworm specimens were reared in a laboratory culture and the *COI* and *Tpi* marker were investigated for their ability to define interstrain hybrids. The potential pitfalls of these markers and the need for new nuclear based markers are highlighted.

2.2 Fall armyworm strains and their importance

Rearing insects for the purpose of meeting human needs such as pest control, animal feed, textiles and research has been taking place in insectaries and laboratories for over 100 years (Cohen 2018). With some records of silkworm cultures dating back to almost 5000 years in China (Vainker 2004, Cohen 2018). The major benefits of rearing insect cultures are that insects can be made available year-round, allowing for their continuous use, and typically very high numbers of insects can be reared as compared to the time and effort to capture them from the field. In terms of research, laboratory reared cultures are also beneficial for the purposes of biological experiments in which physiological aspects such feeding behaviour,

reproduction, life span and life stage duration can be investigated to develop integrated control strategies.

The recent introduction of the fall armyworm, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera, Noctuidae), a notorious corn pest into Africa (2016) and India (2018) has resulted in increased research of this pest (Goergen *et al.* 2016, Kalleshwaraswamy *et al.* 2018). Research investigating control strategies in Africa have already started, with multiple studies using laboratory reared fall armyworm cultures (Sisay *et al.* 2019, Souza *et al.* 2019, Laminou *et al.* 2020, Sokame *et al.* 2021). The use of laboratory reared cultures of this pest is most likely to become more common as the pest continues to spread and its presence intensifies in the various corn growing regions of Africa.

The fall armyworm is represented by two strains, a corn strain which has been found to prefer corn and sorghum and a rice strain which has been found to prefer rice and forage grasses (Pashley *et al.* 1985, Pashley 1986). These two strains were initially identified and defined by differing allelic frequencies on polymorphic loci revealed by allozyme electrophoresis analysis of samples collected in Mexico, Southeastern United States and the Caribbean (Pashley *et al.* 1985, Pashley 1986). These strains have been shown to exhibit differences in fitness on different host-plants (Pashley 1988). A strong relationship between strain identity and host preference has been confirmed by screening multiple fall armyworm populations collected from different host plants, even though in a few instances some of the specimen strain identities did not match their expected host preference (Murúa *et al.* 2015). As well as differences in host preference, significant differences in developmental performance were observed when larvae from either of the strains fed on different Bermuda grass cultivars (Pashley *et al.* 1987). These differences in host preference and developmental performance on hosts act as potential reproductive barriers resulting in larger corn or rice populations and smaller hybrid populations of the fall armyworm (Murúa *et al.* 2015).

In addition to habitat isolation, two other reproductive isolation mechanisms, which limit hybridisation of the strains are differences in pheromone compositions released by adult female moths (Groot *et al.* 2008) and differences in mating times at night (Pashley *et al.* 1992). Strains have even been observed to peak in their larval densities during different seasons, the corn strain in spring and midsummer and the rice strain in late summer and autumn (Pashley *et al.* 1992). Despite these reproductive barriers, hybridisation between the two strains occurs and strains may occur on host plants that do not match their strain identity (Murúa *et al.* 2015).

2.3 Genome comparison of the strains

The genomes of these two strains were sequenced and compared in study by Gouin *et al.* (2017) who found significant differences in the nuclear and mitochondrial genomes of the strains. The specimens used for genome sequencing were laboratory reared, two corn strain male larvae and one rice strain larva. The strain identities were determined using *FR* (fall armyworm Rice strain) repeat markers (Nagoshi and Meagher 2003). The polyphagous nature of the pest was evident in its large array of chemosensory genes, allowing it to identify a wide variety of suitable hosts, yet between the two strains there was no significant difference in number of genes. Genes related to enzyme families responsible for detoxification were also investigated, namely cytochrome P450's (CYPs), glutathione-S-transferases (GSTs), esterases (CCEs) and UDP-glycotransferases (UGTs). Detoxification genes, which were strain specific, were identified for each of the above-mentioned gene families. They showed that these differences in detoxification genes could potentially serve as strain-specific markers, but there have been no further publications that have followed on from this. These significant differences in detoxification genes and additionally, a number of digestion related genes, may be the reason for varying adaption of the specific strains to their preferred host plants (Gouin *et al.* 2017). These genetic differences may also be the reason for differences in susceptibility towards different plant cultivars and certain pesticides (Pashley and Martin

1987, Pashley *et al.* 1987, Nagoshi and Meagher 2004). Strain identification is therefore needed to manage this pest effectively and to conduct research which is replicable.

2.4 Molecular identification of Fall armyworm strains

The fall armyworm strains are morphologically indistinguishable and for this reason molecular methods are relied upon to accurately differentiate and identify the strain identity of specimens. Since the discovery of the two strains by allozyme electrophoresis, differences between the strains in nuclear and mtDNA restriction fragment length polymorphisms (RFLP), nuclear amplified fragment length polymorphisms (AFLP), nuclear tandem repeat sequences, restriction sites in specific mitochondrial genes and single nucleotide polymorphisms (SNP) in sex-linked genes have been identified and used as markers for strain determination (Lu *et al.* 1992, Lu *et al.* 1994, Lu and Adang 1996, McMichael and Prowell 1999, Levy *et al.* 2002, Nagoshi 2010) (Table 2.1).

Table 2.1: Major techniques in defining and identifying fall armyworm strains.

Methods	How strains were defined	autosomal DNA / sex-linked / mtDNA	Title	Authors	Year
Comparison of allelic frequencies on polymorphic loci by allozyme electrophoresis	Host plants specimens were collected from and significant differences in 5 polymorphic loci	Autosomal	Genetic population structure of migratory moths: the Fall Armyworm (Lepidoptera: Noctuidae)	Dorothy P. Pashley, Seth J. Johnson, Alton N. Sparks	1985
			Host-associated genetic differentiation in Fall Armyworm (Lepidoptera: Noctuidae): a sibling species complex?	Dorothy P. Pashley	1986
Restriction fragment length polymorphism (RFLP) analysis. (variation in DNA sequences recognised by restriction enzymes)	RFLP variation in nuclear DNA, most significant variation using markers H2/*Eco*RI and H4/*Eco*RV (probe and restriction site)	Autosomal	RFLP analysis of genetic variation in Northern American populations of the fall armyworm moth *Spodoptera frugiperda* (Lepidoptera: Noctuidae)	Yang-Jiang Lu, M. J. Adang, D. J. Isenhour and G. D. Kochert	1992
Southern blot analysis of *FR* (*Found in rice strain*) tandem repeat sequences sequences digested using *Hae*III restriction endonuclease	Multiple *FR* repeats distinguished rice strain from corn strain	Sex-linked	Molecular characterization of a strain-specific repeated DNA sequence in the fall armyworm Spodoptera frugiperda (Lepidoptera: Noctuidae)	Yang-Jiang Lu, G. D. Kochert, D. J. Isenhour and M. J. Adang	1994
Restriction digestion of mitochondiral DNA using restriction endonucleases *Hae*III and *Msp*I	Significantly different restriction banding patterns differentiated the strains	mtDNA	Distinguishing fall armyworm (Lepidoptera: Noctuidae) strains using a diagnostic mitochondrial DNA marker	Yang-Jiang Lu and M. J. Adang	1996
Amplified fragment length polymorphism (AFLP) analysis	Twenty-eight multilocus AFLP fingerprints seperated specimens into the two strains	Autosomal	Differences in Amplified Fragment-Length Polymorphisms in Fall Armyworm (Lepidoptera: Noctuidae) Host Strains	Margaret McMichael and Dorothy P. Pashley	1999
PCR amplification of *COI* gene and restriction digestion using *Msp*I	Different banding patterns distinguish the two strains (*Msp*I site only on corn strain gene)	mtDNA	Strain Identification of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) Insects and Cell Line: PCR-RFLP of Cytochrome Oxidase C Subunit I Gene	Hazel C. Levy, Alejandra Garcia-Maruniak and James E. Maruniak	2002
Sequencing of the *Tpi* gene	Single nucleotide polymorphisms differentiated the two strains	Sex-linked	The Fall Armyworm triose phosphate isomerase (Tpi) gene as a marker of strain identity and interstrain mating	Rodney N. Nagoshi	2010

Although many of these strain diagnostic techniques are not used in more recent publications (last 10 years), they are useful in understanding how these strains have been understood and defined in the past. Currently two genetic markers are frequently used when confirming fall armyworm identity at species and strain identity level. These genetic markers involve amplifying different portions of the mitochondrial gene, *cytochrome oxidase I* (*COI*), and the sex-linked nuclear gene, *triosephosphate isomerase* (*Tpi*) (Nagoshi 2010, 2012). The *COI* marker has been greatly beneficial in strain identification over the years. Strain identification techniques before the *COI* marker required several hundred nanograms of mtDNA and labelled probes (Lu and Adang 1996), the first PCR-RFLP *COI* method only needed a few nanograms and could be performed directly from total DNA extracts (Levy *et al.* 2002). Strain haplotypes have been defined by sequencing different *COI* gene regions (Nagoshi 2012, Nagoshi *et al.* 2012). Strain haplotype determination based on *COI* gene regions have allowed for the investigation of fall armyworm migration patterns and the origins of fall armyworm populations, which have recently been introduced into Africa and India (Nagoshi 2012, Nagoshi *et al.* 2019b).

The more recently developed *Tpi* marker, has also greatly added to understanding fall armyworm strain identities (Nagoshi 2010). The *Tpi* marker has allowed for the identification of male interstrain hybrids and when compared to the *COI* marker was found to be more reliable, in terms of correlating with the specimens' host plant preference (Nagoshi 2012). This marker has been more exclusively used in studying population structure and migratory patterns of the fall armyworm in which fall armyworm male moths are captured through the use of pheromone traps (Nagoshi *et al.* 2018, Nagoshi *et al.* 2019a).

These markers however are not perfect and have some notable issues. The mitochondrial gene, *COI*, is maternally inherited as part of the mitochondrial genome. This means that using this *COI* marker in isolation only provides information pertaining to the maternal parent, while information regarding the paternal strain identity is unaccounted for. For example, a fall

armyworm larvae for which the *COI* strain identity has been confirmed to be rice strain may have either a rice strain mother and rice strain father or a rice strain mother and a corn strain father. Hybridisation of the two strains is therefore not detected when only using the *COI* marker. The *Tpi* marker when used in isolation is only beneficial when the sexes of the specimens are known and most beneficial when the specimens are male. This marker, is based on a sex-linked chromosome present as a single copy in females and as two copies in males. This means that this marker can only be used to identify hybrids in male specimens where heterozygotes can be scored, in contrast to females that are hemizygous at this locus (Nagoshi *et al.* 2017). Using both genetic markers seems to be the best way forward, although the *Tpi* gene may only give insight into hybridisation for male specimens, the use of both genetic markers have the potential for being discordant (markers indicating opposing strains) which may indicate hybridisation in the specimen's ancestry (Nagoshi *et al.* 2017).

2.5 Strain typing of the fall armyworm in Africa

In Africa, the presence of this pest and its strains have primarily been confirmed through sequencing portions of the *COI* gene (Goergen *et al.* 2016, Cock *et al.* 2017, Nagoshi *et al.* 2017, Jacobs *et al.* 2018). The strong flight capabilities of the adult moths, the prolificacy of the adult moths, and polyphagous nature of the larvae (Cock *et al.* 2017) have all contributed to the rapid spread of the pest on the African continent. Caterpillar specimens collected from Togo, West Africa, were screened to determine their strain identities using both *COI* and *Tpi* markers (Nagoshi *et al.* 2017). These caterpillars were collected from corn, and from these 65% were corn strain *COI* and the remaining 35% were rice strain *COI*. When using the *Tpi* marker, 94% were corn strain *Tpi*, only 3% were rice strain *Tpi* and the remaining 3% were *Tpi* hybrids. Only 60% of the specimens were concordant for both markers for corn strain, while 34% of the specimens were discordant, having a *COI* rice strain and a *Tpi* corn strain (Nagoshi *et al.* 2017). Only 2% were discordant with *COI* corn strain and *Tpi* rice strain, and

the *Tpi* hybrid specimens were all *COI* corn strain. These results demonstrated that the corn strain was the dominant strain causing damage to corn in Togo, while the rice strain makes up a lower percentage of the fall armyworm population affecting corn in the region. These results also illustrate that had one marker been used in isolation the strain identities of the fall armyworm populations would not have been accurately determined (Nagoshi *et al.* 2017).

Further investigation into the specific strain haplotypes present in the various fall armyworm populations in Africa revealed that the majority of these populations comprised of only two *COI* corn strain haplotypes and one *COI* rice strain haplotype (Nagoshi *et al.* 2019a). This revealed the fall armyworm populations present in Africa to be less genetically diverse than the populations present in the Western Hemisphere from which this pest originates. Although there are not many different strain haplotypes present in the Africa, there is still a high level of complexity when trying to understand the fall armyworm populations present.

The initial fall armyworm populations evaluated in African countries such as Togo where determined to be majority corn strain, but when evaluating the fall armyworm populations in 11 different countries in Sub-Saharan Africa it was discovered about 80% of the population has been derived from interstrain hybridisation (Nagoshi *et al.* 2019a). The pure corn strain fall armyworm (determined to be corn strain for both markers) was a minority of the population and the pure rice strain fall armyworm was effectively non-existent. The reason for this population composition could be that the initial invasive specimens were interstrain hybrids. The biological characteristics of the interstrain hybrids are unknown and this then raises the questions of how effective pest management should be approached as well as how to correlate research done with these populations to the broader body of fall armyworm research in which hybrids are the minority and, in many cases, unaccounted for. Further research needs to be done to characterise these hybrids and more accurate genetic markers need to be developed (Nagoshi 2019).

2.6 Laboratory reared culture and the strain issue

In the current study, a fall armyworm culture was reared and maintained in order to perform various bioassays to investigate the role of corn secondary metabolites against fall armyworm larvae. The fall armyworm culture available at the University of Pretoria, Zoology and Entomology department, was originally established with larvae collected from corn at the university's experimental farm in Hillcrest, Pretoria. This culture was defined as *COI* corn strain. Another fall armyworm culture was reared at the University of Pretoria, *COI* rice strain fall armyworm which were acquired from the Agricultural Research Council (ARC) – Plant Health and Protection Institute fall armyworm culture kept at Roodeplaat (collected from sites all across RSA, predominantly from corn). These two cultures most likely mixed as they were reared in the same laboratory. The culture was supplemented once (June 2019) and restarted (June 2020) with pupae from the ARC culture. An overview of the fall armyworm culture and when samples were collected from the culture for strain determination is illustrated below in Figure 2.1.

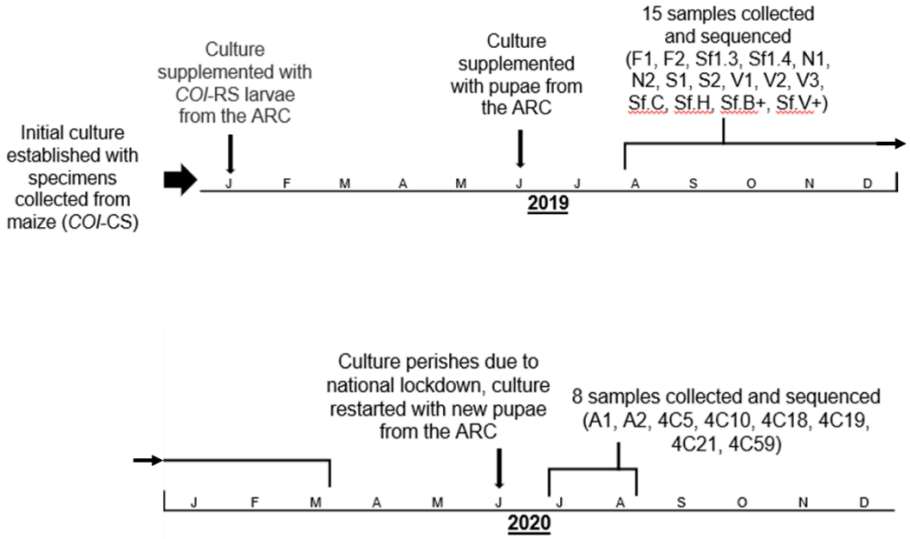


Figure 2.1: Timeline of the fall armyworm culture including when the culture was supplemented with additional specimens and when samples were collected for molecular screening to determine strain identity. The letters on the line represent months of the year, with all the months for 2019 and 2020 being displayed. The codes (e.g. F1) represent individual samples collected, these codes correlate with samples present in table 2.1. The ARC refers to the Agricultural Research Council – Plant Health and Protection Institute based in Roodeplaat, Pretoria.

The fall armyworm were reared in glass cages containing a vermiculite base, wet tissue paper in petri dishes to maintain humidity and wheat plants (Duzi cultivar) for adults moths, eggs and neonates or artificial medium (Kfir 1994) for first to sixth instar larvae as the food source. Petri dishes holding cotton wool soaked with a weak honey solution were included into cages containing wheat plants with adult moths, this served as an additional food source to the adult moths. These cages were kept in IncoCool (Labotec, Midrand, South Africa) and Labcon growth chambers (Labex, Johannesburg, South Africa) which remained at a constant 25°C. The fluorescent lights within these growth chambers emit a light intensity of approximately 10.25 µmol / s / m^2 (760 lux, 12h:12h, L:D cycle). Light intensity was determined using a Heavy Duty Light Meter (Extech®, Massachusetts, United States). The artificial media was regularly replaced and once caterpillars reached the sixth instar stage and started pupating, pupae were collected and transferred to new cages. New cages contained young wheat plants which served as oviposition sites for the adult moths. Eggs would take approximately two days to hatch and the first instar larvae would then be transferred to new cages containing artificial medium. The process from neonate to reproducing adult moth took approximately four weeks.

The cultures were screened using the molecular markers *COI* and *Tpi* and the same methods followed as described in Nagoshi *et. al.* (2017) to determine strain identity. DNA extraction from fall armyworm samples was performed using two different methods. In the first method, individual larval specimens were homogenized in 1.5 ml of phosphate buffered saline (10mM PO43-, 137mM NaCl, 2.7mM KCl) using a mortar and pestle. The homogenate was transferred to a new 2 ml Eppendorf tube. The homogenate was then centrifuged at 6000 x g for 5 min. The supernatant was discarded and the pellet resuspended in 800 µL Genomic Lysis buffer (Zymo Research, Orange, CA, USA). The samples were then incubated at 55°C for 30 min and centrifuged at 6708 x g for 3 min. The supernatant was transferred to Zymo-Spin III-F columns (Zymo Research, Orange, CA, USA) and centrifuged at 8000 x g for 1 min. A total of 500 units of DNA free RNAse A (Norgen Biotek Corp., Ontario, Canada) was added

to each sample and incubated at room temperature for 30 min. An equal amount of phenol:chloroform (1:1) was added, the samples were then inverted and centrifuged at 8117 x g for 5 min. The upper phases of the samples were transferred to new 2 ml Eppendorf tubes and equal amounts of chloroform:isoamyl alcohol (24:1) solution was added. The samples were vortexed, incubated at room temperature for 30 min and were centrifuged at 9660 x g for 10 min. The aqueous phase of the samples was transferred into new 2 ml Eppendorf tubes. Absolute ethanol was added to each sample, mixed and centrifuged at 8117 x g for 15 min. The supernatant was discarded, 70 % ethanol was added and centrifuged at 8117 x g for 10 min. The supernatant was discarded and the sample dried, once dried the samples were resuspended using 100 µL ddH2O. The second method involved using an insect DNA extraction kit, Quick-DNA™ Tissue/Insect Microprep Kit (Zymo Research, Orange, CA, USA). The extraction kit was the primary means of extraction fall armyworm DNA. The manufacturers guidelines were followed. The quality of all DNA samples was evaluated on a NanoDrop™ 2000 Spectrophotometer (Thermo Scientific™, Massachusetts, United States).

Strain identification was performed by PCR amplifying and sequencing the mitochondrial marker, Cytochrome oxidase I (COI), and the sex-lined marker, triosephosphate isomerase (Tpi). To amplify the COX1 gene region the forward primer COI-101F (5'-TTCGAGCTGAATTAGGGACTC -3') and the reverse primer COI-911R (5'-GATGTAAAATATGCTCGTGT -3') were used which produced a 811 bp fragment (Nagoshi *et al.* 2017). To amplify the *Tpi* gene region, the forward primer Tpi-282F (5'-GGTGAAATCTCCCCTGCTATG -3') and the reverse primer Tpi-850R (5'-AATTTTATTACCTGCTGTGG -3') were used, which produced a variable length fragment of about 500 bp (Nagoshi *et al.* 2017). PCR amplification was performed in a 12 µL reaction mixture containing 6.8 µL ddH_2O, 0.6 µL forward primer (10 µM), 0.6 µL reverse primer (10 µM), 3 µL Taq 2X Master Mix Red (Amplicon Inc., Odense, Denmark) and +/- 50 ng genomic DNA. The thermocycle cycling conditions were the same for both primer sets and were as follows: 94°C (1 min), followed by 33 cycles of 92°C (30 s), 56°C (45 s), 72°C (45 s), and a

final segment of 72°C (3 min). The primers and the cycling conditions were obtained from Nagoshi *et al.* 2017. The PCR products were visualised by gel electrophoresis on a 1% agarose gel (80 V, 400 mA, 50 min) with a 100 bp molecular marker (New England Biolabs).

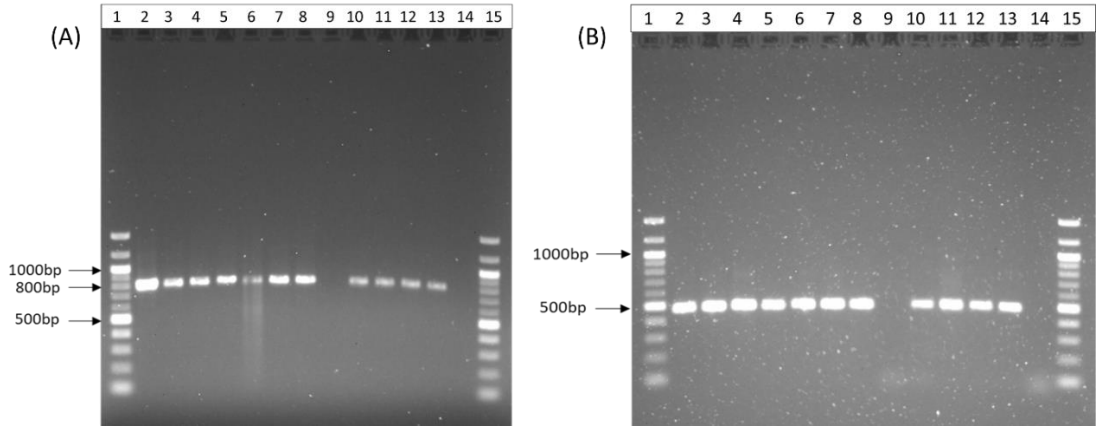
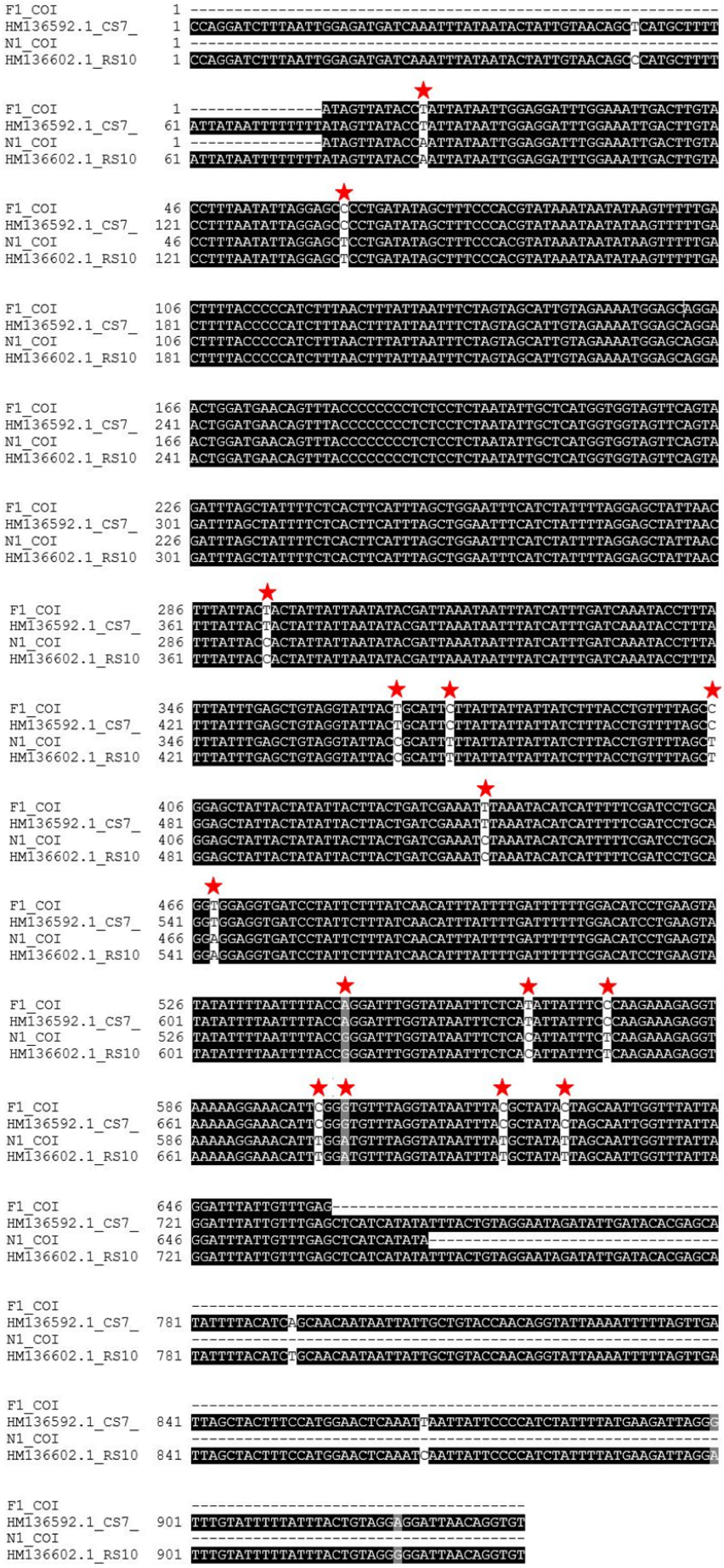


Figure 2.2: *Cytochrome oxidase I* (*COI*) (A) and *Triose phosphate isomerase* (*Tpi*) (B) gene PCR amplification of fall armyworm pupae and larvae gDNA sampled from the University of Pretoria culture. Lanes 2 – 8: fall armyworm pupae samples; Lane 9: non-template/water control; Lanes 10 – 13: fall armyworm larvae samples; Lane 14: non-template/water control; Lane 1 & 15: 100 bp DNA Ladder (New England Biolabs). PCR products were resolved on a 1% TAE ethidium bromide agarose gel subjected to 80 V for 50 min.

Sanger sequencing was performed using the PCR products of the *COI* and *Tpi* gene regions. PCR products were cleaned using ExoSAP-IT (Thermo Fischer Scientific™, Massachusetts, United States). Exosap was added to each of the PCR product samples and the following cycling conditions were utilised: 37°C (15 min) and 80°C (15 min). The sanger sequencing reactions were performed in a 10 µL reaction mixture containing 2 µL BigDye® (Thermo Fischer Scientific™, Massachusetts, United States), 1 µL 5X sequencing buffer, 0.5 µL forward/reverse primer (10 µM), 2.5 µL purified PCR product and 4 µL ddH_2O. The cycling conditions utilised were as follows: 94°C (1 min), followed by 25 cycles of 96°C (10 s), 56°C (5 s) and 60°C (4 min). The samples were then sequenced by the sequencing facility of the University of Pretoria. The *COI* sequences were then used to perform a BLAST query of the NCBI database, multiple SNP differences between the *COI* corn and rice strain genes allowing for strain identification (Figure 2.2 A and B). The *Tpi* sequences were evaluated for SNPs at nucleotide position 183. A mapping strand was generated and aligned to sample sequences to identify the diagnostic SNP position (Figure 2.2 C) (Nagoshi *et al.* 2017, Nagoshi 2019).

(A)



(B)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Spodoptera frugiperda haplotype 2 cytochrome oxidase subunit I (COI) gene, partial cd...	Spodoptera...	1221	1221	100%	0.0	100.00%	777	MN075830.1
Spodoptera frugiperda mitochondrion, complete genome	Spodoptera...	1221	1221	100%	0.0	100.00%	15356	MT732476.1
Spodoptera frugiperda isolate sfC cytochrome c oxidase subunit 1 (COX1) gene, partia...	Spodoptera...	1221	1221	100%	0.0	100.00%	1536	MW665994.1
Spodoptera frugiperda isolate NT-2 cytochrome c oxidase subunit I (COX1) gene, parti...	Spodoptera...	1221	1221	100%	0.0	100.00%	772	MW454866.1
Spodoptera frugiperda isolate GZLD16 cytochrome oxidase subunit I gene, partial cds;...	Spodoptera...	1221	1221	100%	0.0	100.00%	709	MT767440.1
Spodoptera frugiperda isolate CSCXY2 cytochrome oxidase subunit I gene, partial cds...	Spodoptera...	1221	1221	100%	0.0	100.00%	709	MT767439.1
Spodoptera frugiperda haplotype CS7 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1221	1221	100%	0.0	100.00%	938	HM136592.1
Spodoptera frugiperda haplotype CS6 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1221	1221	100%	0.0	100.00%	938	HM136591.1
Spodoptera frugiperda haplotype CS1 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1221	1221	100%	0.0	100.00%	938	HM136586.1
Spodoptera frugiperda haplotype CS5 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1216	1216	100%	0.0	99.85%	938	HM136590.1
Spodoptera frugiperda haplotype CS4 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1216	1216	100%	0.0	99.85%	938	HM136589.1
Spodoptera frugiperda haplotype CS3 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1216	1216	100%	0.0	99.85%	938	HM136588.1
Spodoptera frugiperda haplotype CS2 cytochrome oxidase subunit I (COI) gene, partia...	Spodoptera...	1216	1216	100%	0.0	99.85%	938	HM136587.1

(C)

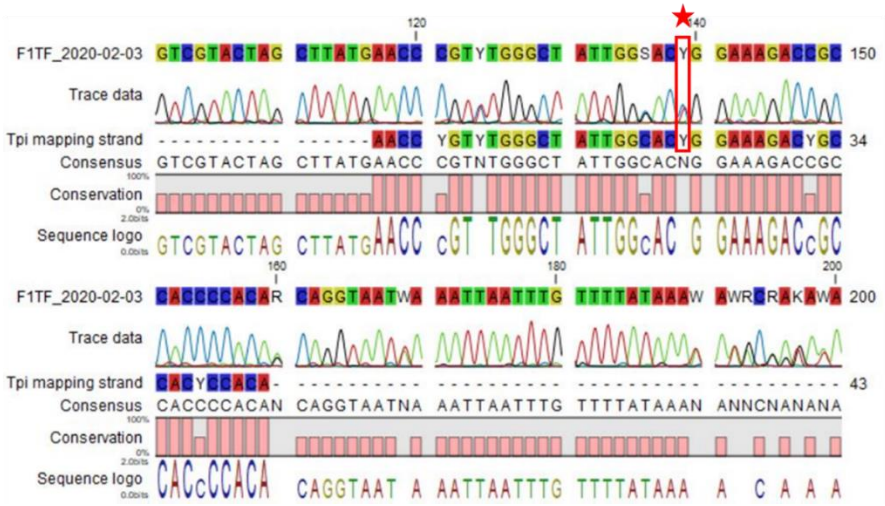


Figure 2.3: Sequencing the *COI* and *Tpi* genes for strain determination. (A) Multiple sequence alignment of the *COI* sequences of samples F1, N1, *S. frugiperda* haplotype C7 and *S. frugiperda* haplotype R10. Diagnostic SNPs are indicated with red stars. (B) BLASTn result for the *COI* gene of sample F1. The top hits matched with the corn strain for this specimen. (B) Alignment of the *Tpi* sequence of sample F1 to the *Tpi* mapping strand to find the diagnostic SNP at base 139 (indicated with a red star and box). A “Y” base call indicating that the specimen is a hybrid strain.

Table 2.2: Summary of the strain identities of fall armyworm samples collected across multiple generations in the laboratory reared cultures.

Culture number	Generation	Sample ID	Sample type*	Date collected	*COI* strain	*Tpi* strain
1	1	F1	Caterpillar (M)	05/08/2019	Corn	Hybrid
	1	F2	Caterpillar (M)	05/08/2019	Corn	Hybrid
	2	Sf1.3	Caterpillar	09/2019	Rice	Corn
	2	Sf1.4	Caterpillar	09/2019	Rice	Corn
	6	N1	Pupae (M)	24/01/2020	Rice	Corn
	6	N2	Pupae (M)	24/01/2020	Rice	Corn
	6	S1	Pupae (M)	24/01/2020	Rice	Corn
	6	S2	Pupae (M)	24/01/2020	Rice	Corn
	6	V1	Pupae (M)	24/01/2020	Rice	Corn
	6	V2	Pupae (M)	24/01/2020	Rice	Corn
	6	V3	Pupae (M)	24/01/2020	Rice	Corn
	8	Sf.C	Caterpillar	16/03/2020	Rice	Corn
	8	Sf.H	Caterpillar	16/03/2020	Rice	Corn
	8	Sf.B+	Caterpillar	16/03/2020	Rice	Corn
	8	Sf.V+	Caterpillar	16/03/2020	Rice	Corn
2	1	A1	Pupae (M)	29/07/2020	Corn	Corn
	1	A2	Pupae (M)	29/07/2020	Corn	Corn
	2	4C5	Caterpillar (M)	30/08/2020	Corn	Hybrid
	2	4C10	Caterpillar	30/08/2020	Corn	Rice
	2	4C18	Caterpillar	30/08/2020	Corn	Rice
	2	4C19	Caterpillar (M)	30/08/2020	Corn	Hybrid
	2	4C21	Caterpillar	30/08/2020	Corn	Rice
	2	4C59	Caterpillar	30/08/2020	Corn	Corn

**Samples denoted with a "(M)" were determined to be male by morphological identification or by analysing the Tpi result.*

In the first fall armyworm culture the strain identity of two specimens in the first generation (F1 and F2) were determined to be *COI* corn strain and *Tpi* hybrid. Samples from the second generation of the same culture (Sf1.3 and Sf1.4) were determined to be *COI* rice strain and *Tpi* corn strain (Table 2.1). It is important to note however that because the sexes of these two specimens (Sf1.3 and Sf1.4) are unknown, they could therefore be female interstrain hybrids displaying hemizygosity at the *Tpi* marker. The strain identity remains constantly *COI* rice strain and *Tpi* corn strain for the samples taken from generation two to eight. The change in strain identity from the first two samples can most likely be attributed to strain associated mating bias. Interstrain mating bias has been observed in laboratory reared cultures and field

collected specimens between *COI* corn strain males and *COI* rice strain females, with *COI* corn strain females and *COI* rice strains males interstrain mating being rare (Nagoshi and Meagher 2003, Nagoshi *et al.* 2006, Nagoshi 2010). The presence of *COI* rice strain specimens in generations two through to eight, indicates that the founder population contained *COI* rice strain females.

In the second fall armyworm culture, the fall armyworm samples, which were part of the first generation (A1 and A2) were determined to be *COI* corn strain and *Tpi* corn strain. This is what some have referred to as pure corn strain specimens (Nagoshi *et al.* 2019b). Fall armyworm samples from the second generation (specifically, 4C5 and 4C19) were determined to be *COI* corn strain and *Tpi* hybrid, indicating the presence of *Tpi* rice strain individuals in the founding population. The rest of the second generation specimens (specifically, 4C10, 4C18, 4C21 and 4C59) were determined to be *COI* corn strain and either *Tpi* rice strain or *Tpi* corn strain (Table 2.1). It is important to note, however, that the sexes for these samples were not known, and therefore these sample could be female interstrain hybrids hemizygous for the *Tpi* marker.

An alternative approach to understanding the genetic composition of the cultures was also used. In this approach specific detoxification genes shown to be exclusive to either strain were screened for, these were a cytochrome P450 (CYP, CYP340L10), UDP-glycosyltransferase (UGT, UGT40-06) and a glutathione S-transferase gene (GST, GST8). This approach was taken from the supplementary information from the Gouin *et al.* (2017) study. In the referenced study, the CYP gene was found to be exclusive to rice strain specimens (*FR* repeats), the UGT gene was found to be exclusive to corn strain specimens (*FR* repeats) and the GST gene was present in both and served as a positive control. The primers used to amplify the CYP gene was the forward primer SfCYP340L10F (5'- GAAGTACGCCATGATGACCTTG -3') and the reverse primer SfCYP340L10R (5'- CCATCAAACATACTCGATCTG -3') resulting in a 528 bp product; to amplify the UGT gene was the forward primer SfUGT40-06F1 (5'-

GGCCATGCCTCGATTTTTCG -3') and the reverse primer SfUGT40-06R (5'-ACTGATTCTTGTAGTCTCGTCCA -3') resulting in a 788 bp product; to amplify the UGT gene was the forward primer SfGST8F (5'- TTGAAGGCATGTGGGGCTC -3') and the reverse primer SfGST8R (5'- TCGAGAAAGTGGAAATGTCAATTT -3') resulting in a 150 bp product (Gouin *et al.* 2017). The same PCR amplification protocol and thermocycling conditions used to amplify the *COI* and *Tpi* gene were used. The PCR products were visualised by gel electrophoresis on a 1% TAE agarose gel (80 V, 400 mA, 50 min) with a 100 bp molecular marker (New England Biolabs).

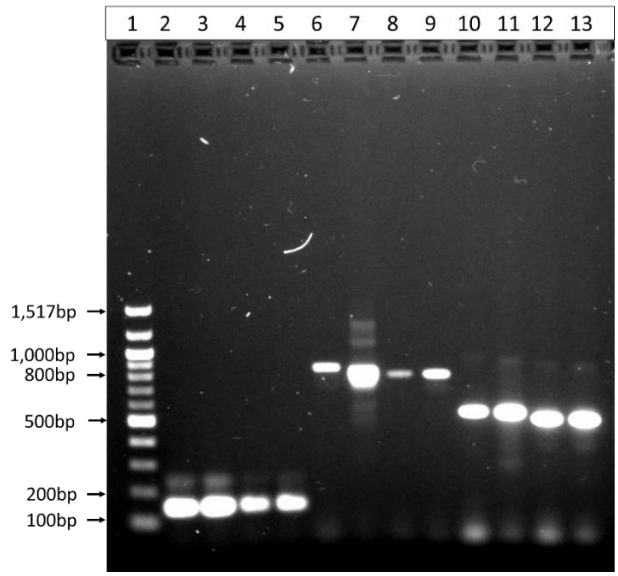


Figure 2.4: Detoxification gene amplification of fall armyworm specimens F1, F2, Sf1.3 and Sf1.4. Lanes 2 – 5: GST gene amplification; Lanes 6 – 9: UGT gene amplification; Lanes 10 - non-template/water control; Lanes 10 – 13: fall armyworm larvae samples; Lane 1: 100 bp DNA Ladder (New England Biolabs).

The CYP (rice strain exclusive) and UGT (corn strain exclusive) genes were amplified from all four samples. This shows that various detoxification genes, which are significantly different between the two strains are being exchanged as intrastrain hybridisation takes place. Hybridisation over multiple generations could result in a more homogenous set of detoxification genes within the culture, meaning that specimens from the first two generations could differ significantly from seventh and eighth generation specimens. When using only the *COI* and *Tpi* markers the exchange of these specific detoxification genes would be unaccounted for. The combination of various detoxification genes may result in different

genetic compositions in the laboratory culture. This may vary significantly when compared to field populations. The inability to determine or measure these potentially significant genes using the conventional markers is worrying because it suggests that experiments and pest management strategies based on interstrain hybrid, laboratory reared cultures may not be replicable or effective when using specimens from either the field or even other laboratories. There is therefore a need to develop reliable nuclear markers able to determine the extent of interstrain hybridisation and to correctly identify corn and rice strain specimens. Without this there is the potential risk of producing research, which is difficult to replicate and difficult to link to the larger body of work done on the fall armyworm.

2.7 A Proposed way forward

The ideal way forward is to develop new nuclear based markers to define the strains. These nuclear based markers could include nuclear genes involved in host preference, namely detoxification, gustatory or chemoreceptor genes (Gouin *et al.* 2017). These markers would then be used in conjunction with the mitochondrial *COI* marker and the sex-linked *Tpi* marker. To define corn strain, rice strain and interstrain hybrids, the following table summarises the potential results when using all three markers in combination (Table 2.3).

Table 2.3: Proposed classification of fall armyworm strains.

No.	Male / Female	*COI*	*Tpi*	*Tpi* interpretation	Nuclear marker (W chromosome)	Physiological preference	Is it a hybrid?
1	Male	Corn	Corn Corn	Corn	n/a	Corn-preferred	No
2	Female	Corn	Corn -	inconclusive	Corn	Corn-preferred	No
3	Male	Rice	Rice Rice	Rice	n/a	Rice - preferred	No
4	Female	Rice	Rice -	inconclusive	Rice	Rice - preferred	No
5	Male	Rice	Corn Corn	C	n/a	Corn-preferred	Past hybrid
6	Female	Rice	Corn -	inconclusive	Corn	Corn-preferred	Past hybrid
7	Male	Corn	Corn Rice	CR	n/a	Corn-preferred	hybrid
8	Female	Corn	Rice -	inconclusive	Corn	Corn-preferred	hybrid
9	Male	Rice	Corn Rice	RC	n/a	Corn-preferred	hybrid
10	Female	Rice	Rice -	inconclusive	Corn	Corn-preferred	hybrid
11	Male	Corn	Rice Rice	R	n/a	Rice - preferred	Past hybrid
12	Female	Corn	Rice -	inconclusive	Rice	Rice - preferred	Past hybrid
13	Male	Corn	Corn Rice	CR	n/a	Rice - preferred	hybrid
14	Female	Corn	Corn -	inconclusive	Rice	Rice - preferred	hybrid
15	Male	Rice	Corn Rice	RC	n/a	Rice - preferred	hybrid
16	Female	Rice	Corn -	inconclusive	Rice	Rice - preferred	hybrid

In this proposed model of strain classification, a nuclear gene marker based on the W chromosome (present only in female specimens) is needed to fill the gap of knowledge left when using the *Tpi* marker on specimens of unknown sex or that have been confirmed to be female. Samples one – four in Table 2.2 represent pure strain specimens. Pure strain specimens would be defined as male specimens with concordant *COI* and *Tpi* markers or female specimens with concordant *COI*, *Tpi* and nuclear (W chromosome) markers. These pure strain specimens should have a preference or have been collected from their expected host plants. Male specimens which are discordant for the *COI* and *Tpi* markers would be defined as having hybrid ancestry and the *Tpi* result would reveal the specimens host preference. Female specimens disconcordant for the *COI* and *Tpi* markers, but displaying concordance between the *Tpi* and nuclear (W chromosome) marker would be considered to have hybrid ancestry and should show a preference for their expected host. Male specimens which are hybrid for the *Tpi* marker would be defined as interstrain hybrids. Female specimens which display discordance between the *Tpi* and nuclear (W chromosome) marker would be defined as interstain hybrids. This classification system is more stringent as it enables to

determination of female interstrain hybrids and therefore lowers the risk of interstrain hybridisation being unaccounted for in a laboratory reared culture.

Precautions aiming to minimise strain hybridisation and increase the potential repeatability of experiments using laboratory reared cultures have been described in previous studies (Meagher and Nagoshi 2012). Pairing adult moths and keeping pairs separate until egg oviposition and screening of the adult moths using the *COI* and *Tpi* markers would enable one to group egg batches of the same strain and prevent including hybrid or unwanted strains. This strategy has been used in a number of studies but when used with only the *COI* marker it has shown to be insufficient in ensuring replicable experiments (Meagher and Nagoshi 2012). In a study by Meagher and Nagoshi (2012) they showed that they were not able acquire repeatable results when repeating the same bioassays three times over different years, using the same laboratory reared cultures. The authors highlighted the limitations of the *COI* marker and raised the question of how effective this marker was in determining the strain composition of a culture to the extent that experiments using the culture were repeatable. Other possibilities for the lack of repeatability when using laboratory cultures could also be as a result of the artificial rearing process and / or high levels of genetic variety within strains that have been unaccounted for (Meagher and Nagoshi 2012). When establishing a fall armyworm culture, hybridisation in a specimen's ancestry may have resulted in the presence of opposite strain alleles being present. For example if specimens determined to be *COI* rice strain and *Tpi* rice strain were used to establish a culture, but some of these specimens were descendants of interstrain hybrids then they may carry corn strain alleles. Laboratory rearing of this culture would then place selection pressure which could favour these corn strain alleles. An example of a selection pressure that could cause this is artificial medium such as pinto bean diet as a food source, *COI* corn strain larvae have been observed to develop faster than *COI* rice strain individuals on this diet (Meagher and Nagoshi 2012). This may result in a culture that is significantly different after many generations from the initial founder specimens. The process of laboratory rearing therefore requires careful thought of how these processes could place

undesired selections pressures on the cultures in question. Ultimately, this again emphasizes the need to develop new nuclear gene markers to accurately define fall armyworm strain identities.

2.8 References

Cock, M. J., P. K. Beseh, A. G. Buddie, G. Cafá, and J. Crozier. 2017. Molecular methods to detect *Spodoptera frugiperda* in Ghana, and implications for monitoring the spread of invasive species in developing countries. Scientific Reports 7: 1-10. DOI: 10.1038/s41598-017-04238-y

Cohen, A. C. 2018. Ecology of insect rearing systems: a mini-review of insect rearing papers from 1906-2017. Advances in Entomology 6: 86.

Goergen, G., P. L. Kumar, S. B. Sankung, A. Togola, and M. Tamo. 2016. First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae), a new alien invasive pest in West and Central Africa. PloS One 11: e0165632. https://doi.org/10.1371/journal.pone.0165632

Gouin, A., A. Bretaudeau, K. Nam, S. Gimenez, J. M. Aury, B. Duvic*, et al.* 2017. Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. Scientific Reports 7: 11816. https://doi.org/10.1038/s41598-017-10461-4

Groot, A. T., M. Marr, G. Schöfl, S. Lorenz, A. Svatos, and D. G. Heckel. 2008. Host strain specific sex pheromone variation in *Spodoptera frugiperda*. Frontiers in Zoology 5: 1-13. https://doi.org/10.1186/1742-9994-5-20

Jacobs, A., A. V. Vuuren, and I. H. Rong. 2018. Characterisation of the fall armyworm (*Spodoptera frugiperda* J. E. Smith) (Lepidoptera: Noctuidae) from South Africa. African Entomology 26: 45-49.

Kalleshwaraswamy, C., R. Asokan, H. M. Swamy, M. Maruthi, H. Pavithra, K. Hegbe*, et al.* 2018. First report of the fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae), an alien invasive pest on maize in India. Pest Management in Horticultural Ecosystem 24: 23-29.

Kfir, R. 1994. Rearing the cotton bollworm, *Heliothis armigera*. Techniques of Insect Rearing for the Development of Integrated Pest and Vector Management Strategies: 417-419.

Laminou, S. A., M. N. Ba, L. Karimoune, A. Doumma, and R. Muniappan. 2020. Parasitism of locally recruited egg parasitoids of the fall armyworm in Africa. Insects 11: 430. https://doi.org/10.3390/insects11070430

Levy, H. C., A. Garcia-Maruniak, and J. E. Maruniak. 2002. Strain identification of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) insects and cell line: PCR-RFLP of cytochrome oxidase C subunit I gene. Florida Entomologist 85: 186-190.

Lu, Y., and M. J. Adang. 1996. Distinguishing fall armyworm (Lepidoptera: Noctuidae) strains using a diagnostic mitochondrial DNA marker. Florida Entomologist: 48-48.

Lu, Y. J., M. Adang, D. Isenhour, and G. Kochert. 1992. RFLP analysis of genetic variation in North American populations of the fall armyworm moth *Spodoptera frugiperda* (Lepidoptera: Noctuidae). Molecular Ecology 1: 199-208.

Lu, Y. J., G. Kochert, D. Isenhour, and M. Adang. 1994. Molecular characterization of a strain-specific repeated DNA sequence in the fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae). Insect Molecular Biology 3: 123-130.

McMichael, M., and D. P. Prowell. 1999. Differences in amplified fragment-length polymorphisms in fall armyworm (Lepidoptera: Noctuidae) host strains. Annals of the Entomological Society of America 92: 175-181.

Meagher, R. L., and R. N. Nagoshi. 2012. Differential feeding of fall armyworm (Lepidoptera: Noctuidae) host strains on meridic and natural diets. Annals of the Entomological Society of America 105: 462-470.

Murúa, M. G., R. N. Nagoshi, D. A. Dos Santos, M. M. Hay-Roe, R. L. Meagher, and J. Vilardi. 2015. Demonstration using field collections that Argentina fall armyworm populations exhibit strain-specific host plant preferences. Journal of Economic Entomology 108: 2305-2315.

Nagoshi, R. N. 2010. The fall armyworm triose phosphate isomerase (*Tpi*) gene as a marker of strain identity and interstrain mating. Annals of the Entomological Society of America 103: 283-292.

Nagoshi, R. N. 2012. Improvements in the identification of strains facilitate population studies of fall armyworm subgroups. Annals of the Entomological Society of America 105: 351-358.

Nagoshi, R. N. 2019. Evidence that a major subpopulation of fall armyworm found in the Western Hemisphere is rare or absent in Africa, which may limit the range of crops at risk of infestation. PloS One 14: e0208966. https://doi.org/10.1371/journal.pone.0208966

Nagoshi, R. N., and R. Meagher. 2003. Fall armyworm FR sequences map to sex chromosomes and their distribution in the wild indicate limitations in interstrain mating. Insect Molecular Biology 12: 453-458.

Nagoshi, R. N., and R. L. Meagher. 2004. Seasonal distribution of fall armyworm (Lepidoptera: Noctuidae) host strains in agricultural and turf grass habitats. Environmental Entomology 33: 881-889.

Nagoshi, R. N., R. L. Meagher, and M. Hay-Roe. 2012. Inferring the annual migration patterns of fall armyworm (Lepidoptera: Noctuidae) in the United States from mitochondrial haplotypes. Ecology and Evolution 2: 1458-1467.

Nagoshi, R. N., R. L. Meagher, G. Nuessly, and D. G. Hall. 2006. Effects of fall armyworm (Lepidoptera: Noctuidae) interstrain mating in wild populations. Environmental Entomology 35: 561-568.

Nagoshi, R. N., G. Goergen, H. Du Plessis, J. van den Berg, and R. Meagher. 2019a. Genetic comparisons of fall armyworm populations from 11 countries spanning sub-Saharan Africa provide insights into strain composition and migratory behaviors. Scientific Reports 9: 1-11.

Nagoshi, R. N., G. Goergen, K. A. Tounou, K. Agboka, D. Koffi, and R. L. Meagher. 2018. Analysis of strain distribution, migratory potential, and invasion history of fall armyworm populations in northern Sub-Saharan Africa. Scientific Reports 8: 1-10.

Nagoshi, R. N., I. Dhanani, R. Asokan, H. Mahadevaswamy, C. M. Kalleshwaraswamy, and R. L. Meagher. 2019b. Genetic characterization of fall armyworm infesting South Africa and India indicate recent introduction from a common source population. PLoS One 14: e0217755.

Nagoshi, R. N., D. Koffi, K. Agboka, K. A. Tounou, R. Banerjee, J. L. Jurat-Fuentes*, et al.* 2017. Comparative molecular analyses of invasive fall armyworm in Togo reveal strong similarities to populations from the eastern United States and the Greater Antilles. PLoS One 12: e0181982.

Pashley, D. P. 1986. Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): a sibling species complex? Annals of the Entomological Society of America 79: 898-904.

Pashley, D. P. 1988. Quantitative genetics, development, and physiological adaptation in host strains of fall armyworm. Evolution 42: 93-102.

Pashley, D. P., and J. A. Martin. 1987. Reproductive incompatibility between host strains of the fall armyworm (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 80: 731-733.

Pashley, D. P., S. J. Johnson, and A. N. Sparks. 1985. Genetic population structure of migratory moths: the fall armyworm (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 78: 756-762.

Pashley, D. P., S. S. Quisenberry, and T. Jamjanya. 1987. Impact of fall armyworm (Lepidoptera: Noctuidae) host strains on the evaluation of bermuda grass resistance. Journal of Economic Entomology 80: 1127-1130.

Pashley, D. P., A. M. Hammond, and T. N. Hardy. 1992. Reproductive isolating mechanisms in fall armyworm host strains (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 85: 400-405.

Sisay, B., T. Tefera, M. Wakgari, G. Ayalew, and E. Mendesil. 2019. The efficacy of selected synthetic insecticides and botanicals against fall armyworm, *Spodoptera frugiperda*, in maize. Insects 10: 45. DOI: 10.3390/insects10020045

Sokame, B. M., J. Obonyo, E. M. Sammy, S. A. Mohamed, S. Subramanian, D. C. Kilalo, *et al.* 2021. Impact of the exotic fall armyworm on larval parasitoids associated with the lepidopteran maize stemborers in Kenya. Biocontrol 66: 193-204.

Souza, C., L. Silveira, R. Pitta, J. Waquil, E. Pereira, and S. Mendes. 2019. Response of field populations and Cry-resistant strains of fall armyworm to Bt maize hybrids and Bt-based bioinsecticides. Crop Protection 120: 1-6. https://doi.org/10.1016/j.cropro.2019.01.001

Vainker, S. J. 2004. Chinese silk: a cultural history, British Museum Press.

Chapter 3: Bioinformatic analysis of *ent*-copalyl diphosphate synthase 2 (An2) across multiple maize inbred lines

3.1 Introduction

Plants respond to pathogens by synthesizing numerous defence compounds through multiple different metabolic pathways (Harris *et al.* 2005). One of these metabolic pathways is the terpenoid pathway. Terpenoids make up one of the most structurally diverse groups of compounds with at least 40 000 compounds (Tholl 2015). These compounds are all derived from the same two precursor molecules, isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP) (Tholl 2015). Condensation of DMAPP with multiple IPP compounds by prenyl transferases yield a variety of chain lengths – namely one, two or three added IPP compounds produces geranyl diphosphate (GPP), farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP), respectively (Tholl 2015, Block *et al.* 2019). These compounds are then used to produce monoterpenes, diterpenes, triterpenes, tetraterpenes and polyterpenes (Block *et al.* 2019).

Diterpenes make up various plant hormones and secondary metabolites. Gibberellic acids (GA), a diterpenoid group of compounds, are plant hormones that serve as plant growth regulators and stimulate physiological process such as seed germination, the transition of meristem to shoot growth and flowering (Gupta and Chakrabarty 2013). The gibberellin synthesis pathway is therefore predominantly active in young plant tissue such as a developing embryo and in growth points such as the apical meristem (White *et al.* 2000, Nelissen *et al.* 2012). This pathway begins with the precursor molecule GGPP and utilises 3 types of enzymes, namely terpene synthases, cytochrome P450 mono-oxygenases and 2-oxoglutarate dependent dehydrogenases to produce bioactive GA's. Terpene synthase enzymes, *ent*-copalyl diphosphate synthase (generally termed An, An1, *ent*-CPS and *ent*-CPS1) and *ent*-kaurene synthases (*ent*-KS, ZmKSL3 / ZmKSL5 / ZmTPS1), which are located in the plastids, convert GGPP molecules into *ent*-kaurene (Figure 3.1) (Sun and Kamiya 1997, Gupta and Chakrabarty 2013, Ding *et al.* 2019). P450 mono-oxygenases, namely *ent*-kaurene

and *ent*-kaurenoic acid oxidases (situated in the outer membrane of the plastid and the endoplasmic reticulum, respectively), utilise *ent*-kaurene and synthesize GA_{12} (Helliwell *et al.* 2001, Gupta and Chakrabarty 2013). This GA is then further oxidized by other oxidase enzymes to synthesize the larger variety of gibberellic acid's. The *ent*-copalyl diphosphate synthase enzyme present in the GA synthesis pathway is present in all plants (Gupta and Chakrabarty 2013). In some instances these enzymes have undergone gene duplication events and formed part of new secondary metabolite synthesis pathways. Examples of these secondary metabolites are oryzalexins and kauralexins in the genera *Oryza* and *Zea* respectively (Prisic *et al.* 2004, Schmelz *et al.* 2011, Toyomasu *et al.* 2015). The *ent*-copalyl diphosphate synthase proteins are referred to in this study as An1 / An2 when referring to the enzyme in maize and *ent*-CPS / *ent*-CPS1 / *ent*-CPS2 when referring to the enzyme in non-maize genera. Enzymes denoted with the number 1 or without any number are involved in primary metabolism (GA synthesis).

The diterpene synthases *ent*-CPS and *ent*-KS, have been repeatedly duplicated and sub-functionalized to form the kauralexin synthesis pathway (Gao *et al.* 2012, Ding *et al.* 2019). The kauralexin synthesis pathway is responsible for the production of the secondary metabolites, kauralexins. Kauralexins are synthesized in response to various biotic (fungi and insect herbivory) and abiotic (drought) stressors (Harris *et al.* 2005, Schmelz *et al.* 2011, Vaughan *et al.* 2015). Although the kauralexin synthesis pathway produces completely different compounds compared to the GA synthesis pathway the same precursor molecule, GGPP, is used by both pathways (Figure 3.1). The *ent*-copalyl diphosphate synthase enzymes involved in secondary metabolism as denoted with the number 2, for example An2 in maize or *ent*-CPS2 in non-maize genera.

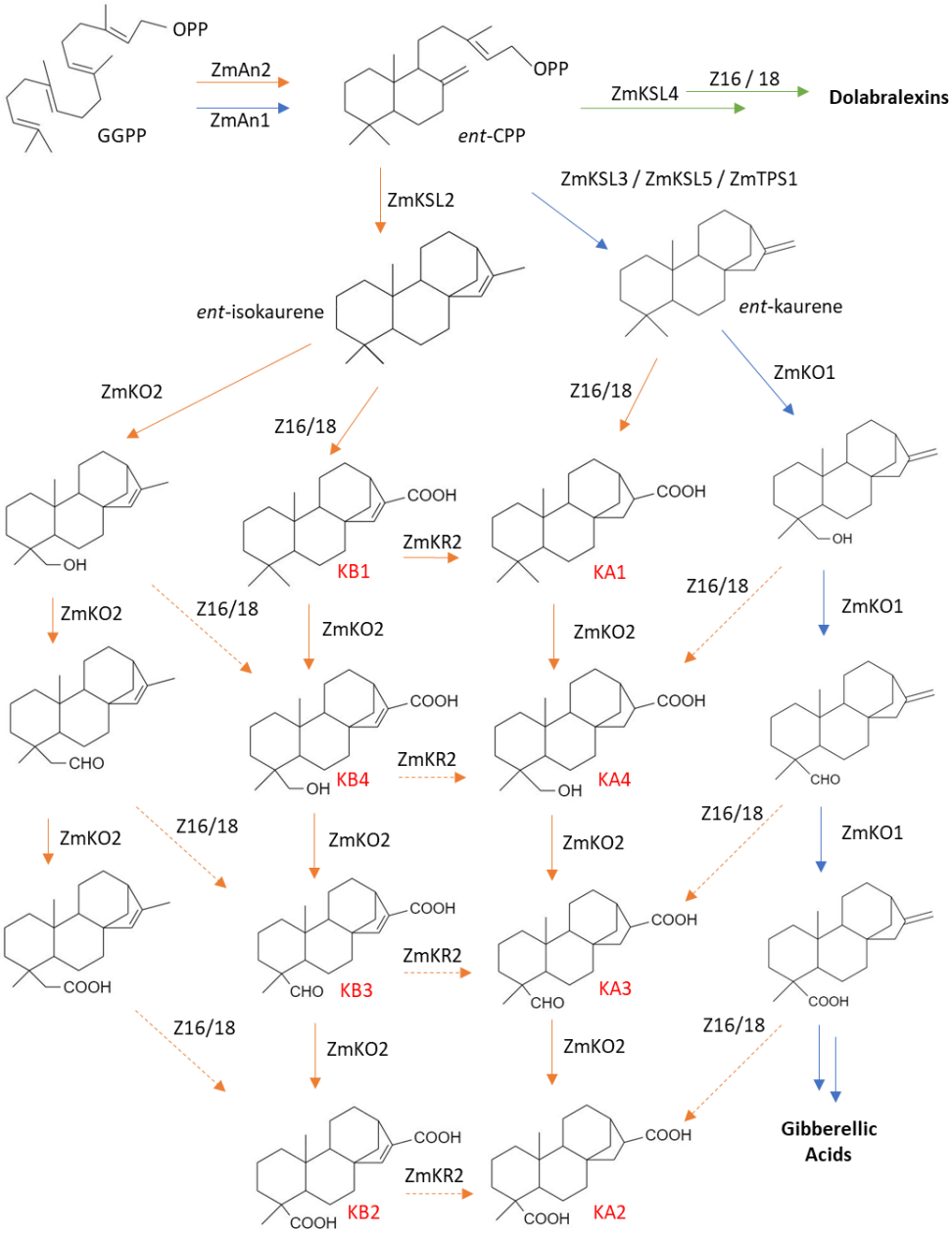


Figure 3.1: Biosynthesis pathway overlap of phytohormone gibberellic acid and defence diterpenoid kauralexin and dolabralexin synthesis pathways. The blue arrows indicate the gibberellic acid phytohormone pathway, the green arrows indicate the plant defence dolabralexin pathway and the orange arrows indicate the kauralexin synthesis pathway. Solid line arrows indicate enzyme catalysis which has been experimentally confirmed, while dashed line arrow indicate predicted catalytic reactions based on enzyme promiscuity. Double arrows indicate omitted steps as means of summarizing the figure. Red text indicates the specific kauralexin compound at the particular step in the biosynthetic pathway. The identity of the enzymes involved in particular steps are indicated in black text either on or next to the arrow, indicating the reaction it catalyses. (Figure adapted from Ding *et al.* 2019)

The cyclization of the universal diterpenoid precursor, GGPP, is catalysed by both Anther ear 1 (ZmAn1) and Anther ear 2 (ZmAn2) *ent*-copalyl diphosphate synthases, forming *ent*-copalyl diphosphate (CDP) (Harris *et al.* 2005, Schmelz *et al.* 2011, Ding *et al.* 2019). Various *ent*-kaurene like synthases (KSL) are able to partition *ent*-copalyl diphosphate into three different synthesis pathways, the hormone pathway of GA synthesis (ZmKSL3 / ZmKSL5 / ZmTPS1) or the secondary metabolite pathways of kauralexin (ZmKSL2) and dolabralexin (ZmKSL4) synthesis. The kauralexin synthesis pathway starts with the *ent*-kaurene like synthase, ZmKSL2, as it utilises CDP to synthesise mainly *ent*-isokaurene (95%) and *ent*-kaurene (5%). *Ent*-isokaurene is not used in GA synthesis, thus separating the kauralexin pathway from the GA synthesis pathway. There are two series of kauralexin compounds, namely kauralexin A's (KA) and B's (KB) which have 4 identified compounds in each (Figure 3.1). These compounds differ in extent of oxidation and the presence (KB) or absence (KA) of a double between carbons C-15 and C-16. *Ent*-isokaurene is oxidised multiple times on the carbons C-17 and C-19 by promiscuous cytochrome P450 monooxygenases, namely Kaurene oxidase 2 (ZmKO2) and ZmCYP71Z16/Z18 (Z16/18). The oxidation reactions of these P450's of *ent*-isokaurene synthesise kauralexin B's – KB1 (*ent*-isokaurene oxidated by Z16/18 to form a carboxyl group on carbon C-19), KB4 (KB1 oxidated by ZmKO2 to form an alcohol group on carbon C-17), KB3 (KB4 oxidated by ZmKO2 to form an aldehyde group on carbon C-17) and KB2 (KB3 oxidated by ZmKO2 to form a carboxyl group on carbon C-17). A steroid 5-α reductase enzyme, kauralexin reductase 2 (ZmKR2), utilises kauralexin B compounds to perform a reduction reaction at carbons C-15 – C16, reducing the double bond to a single bond (Figure 3.1). This reductase is therefore able to synthesize *ent*-kaurene associated defence compounds without the kauralexin pathway needing to synthesize *ent*-kaurene (Ding *et al.* 2019). Mutations in the *ZmAn2* and *ZmKSL2* genes separately results in near complete loss of kauralexin synthesis (Vaughan *et al.* 2015, Christensen *et al.* 2018, Ding *et al.* 2019), thus showing that the kauralexin synthesis pathway although it uses CDP, a precursor also to GA synthesis, is regulated in such a way that co-expression without dysregulation of the hormone pathway is possible (Ding *et al.* 2019). Some of these regulatory methods being the

induction of ZmAn2, ZmKSL2, Z16/18, ZmKO2 and ZmKR2 enzyme synthesis as a response to fungal pathogen invasion, the synthesis and use of *ent*-isokaurene, which is not used in GA synthesis, and *Z16/18* utilising the trace amounts of *ent*-kaurene synthesized by ZmKSL2 (Ding *et al.* 2019).

The kauralexin pathway is present across all the various maize inbred lines, but the levels of total kauralexins synthesized differ between them (Schmelz *et al.* 2011, Ding *et al.* 2019). Total kauralexin levels in the scutella of 10 day old seedlings differed across 19 diverse inbred lines (Schmelz *et al.* 2011). Another study investigating total kauralexin induction due to the presence of *Fusarium* hyphae on maize stems determined that different levels of total kauralexins were induced over 27 maize inbred lines (Ding *et al.* 2019). When comparing these two studies, the results differed for some maize inbred lines. In the study which looked at maize seedlings, the inbred line Ms71 synthesized the highest amount of kauralexins compared to other inbred lines tested whereas in the second study (involving the presence of *Fusarium* hyphae), Ms71 synthesized the lowest amount of kauralexins. Some inbred lines were more consistent over the two studies, for example the maize inbred line CML333 was among the top 4 highest kauralexin synthesizing inbred lines in both studies. The maize inbred line W22, synthesized higher total kauralexin levels when compared to the model inbred line, B73 (Schmelz *et al.* 2011).

The kauralexin synthesis potential across maize inbred lines could be linked to variations in the amino acid sequence and 3D structure of the proteins involved in the synthesis pathway. Variation in protein structure can result in differences in binding affinity to the specific precursor molecules and in turn affect catalytic activity, which would influence the rate and potentially the amount of kauralexins produced by a maize line. Experimental methods of determining protein structure include x-ray crystallography, nuclear magnetic resonance (NMR) and electron microscopy, all of which require costly equipment (Tastan Bishop *et al.* 2008). Alternative methods to determine protein structure, which are more cost effective and

generally more accessible is that of computational biology. Homology modelling is a computational method, which predicts the 3D structure of a protein based on a template protein, a protein, which displays sequence similarity and that has an experimentally determined structure (Tastan Bishop *et al.* 2008). These homology models can in turn be used to investigate the catalytic sites present and binding affinities of these sites to their specific ligands. Although this approach is quicker and more cost effective compared to laboratory tests, it is still predictive and would need to be confirmed experimentally.

The *ent*-copalyl diphosphate enzyme (ZmAn2) is the focus of this study because of its role in maize defences, as it catalyses the first step in the kauralexin synthesis pathway and has been experimentally proven to be critical in the synthesis of kauralexins through gene-knockout mutations (Vaughan *et al.* 2015, Block *et al.* 2019, Ding *et al.* 2019). The aim of this study is to compare the amino acid sequences and protein structures of the *ent*-copalyl diphosphate synthase enzymes (ZmAn2) from multiple maize inbred lines in order to determine whether variation in this enzyme could result in total kauralexin level differences which may affect maize defences. The amino acid sequences will be compared, protein modelling performed and the GGPP ligand will be docked to the protein structures to determine if there are any differences in ligand binding affinity. The ZmAn2 proteins are expected to differ in their binding affinity across different maize inbred lines and potentially give insight into why different maize inbred lines synthesized different total levels of kauralexins.

3.2 Methods

3.2.1 Phylogenetic analysis and comparison of ent-copalyl diphosphate synthases

Phylogenetic analyses were performed on *ent*-copalyl diphosphate synthase 1 (An1 / *ent*-CPS / *ent*-CPS1) and *ent*-copalyl diphosphate synthase 2 (An2 / *ent*-CPS2) amino acid sequences from *Z. mays*, *Setaria italica*, *Oryza sativa Japonia*, *O. sativa Indica, Panicum virgatum*, *Triticum aestivum*, *Arabidopsis lyrata* and *Arabidopsis thaliana*. *Zea mays* An1 and An2 amino acid sequences were obtained by performing a Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990) search of the different maize genomes available on maizeGDB (Lawrence *et al.* 2004, Monaco *et al.* 2013) using the maize inbred line B73 An1 and An2 amino acid sequences as the query sequences. Sequences were obtained from a total of 37 different maize inbred line genomes. The *ent*-copalyl diphosphate synthase amino acid sequences of the other genera were obtained by performing a BLAST search on the National Center for Biotechnology Information (NCBI) database using the maize inbred line B73 An2 amino acid sequence as the query sequence. The dataset (supplementary table 1) was aligned using the online software Multiple Alignment using Fast Fourier Transform (MAFFT) Multiple Sequence Alignment Version 7 (Katoh and Standley 2013) using the default settings. The alignment was then imported into GBlocks version 0.91b online software (Castresana 2000) to curate the alignment, selecting for conserved blocks. MEGAX (Molecular Evolutionary Genetics Analysis, (Kumar *et al.* 2018)) was used to produce a maximum-likelihood phylogenetic tree, which was statistically tested with 1000 bootstrap iterations using the JTT (Jones-Taylor-Thornton) matrix based model (Jones *et al.* 1992) and the default settings. Bootstrap support values lower than 70% were set to not be visible on the final phylogenetic tree (Figure 3.2).

The amino acid sequences of the maize inbred line B73 (both An1 and An2), a non-*Zea* plant genus *ent*-CPS1 and *ent*-CPS2 protein (*O. sativa Japonica*) and the outgroup *ent*-CPS protein (*A. thaliana*) were imported, aligned subject to pairwise comparisons using CLC Main Workbench version 8 (Qiagen, Hilden, Germany). The pairwise comparison focused on amino acid identity and percentage identity across the whole amino acid sequences (Figure 3.3). A direct comparison of the model maize inbred line B73 An1 and An2 amino acid sequences was performed. Both sequences were imported and aligned in CLC Main Workbench version 8 and the sequence alignment was then imported into GBlocks version 0.91b online software to highlight identical, conserved regions. The catalytic sites of the amino acid sequences were identified by submitting the sequences to InterPro Scan (Blum *et al.* 2021). The terpene synthase and metal binding sites were identified and location of these sites on the amino acid sequences noted (Figure 3.4 and 3.5). The terpene synthase catalytic sites were then indicated on the GBlocks alignment to highlight the degree on conservation across the two proteins.

3.2.2 Catalytic site analysis of ZmAn2

The catalytic sites, namely the terpene synthase and metal binding sites, which were identified on maize inbred line B73 An2 amino acid sequence were illustrated in the form of a diagram (Figure 3.5). The identified terpene synthase site was then used to identify the terpene synthase catalytic sites of the *A. thalina ent*-CPS and remaining *Z. mays* An2 amino acid sequences using the online multiple sequence alignment software MAFFT. The alignment was trimmed using CLC Main Workbench version 8. MEGAX was used to produce a maximum-likelihood phylogenetic tree which was statistically tested with 1000 bootstrap iterations using the JTT matrix based model and the default settings. Bootstrap support values lower than 70% were set to not be visible on the final phylogenetic tree. *Zea mays* An2 sequences were selected from each major clade of the phylogenetic tree to use for protein modelling and ligand docking (Figure 3.6). These selected maize inbred lines along with the *Arabidopsis* outgroup,

were then aligned and subject to pairwise comparisons using CLC Main Workbench version 8 in order to investigate amino acid identity and percentage identity across the whole amino acid sequences (Figure 3.8). The catalytic sites for these maize An2 proteins were then aligned in CLC Main Workbench version 8 and the amino acid identities which differed across the maize lines noted.

3.2.3 Protein modelling and ligand docking of ZmAn2

Selected amino acid sequences, namely the *Z. mays* An2 proteins for B73, W22, CO387, CML333, Tx303, PE0075 and the *A. thaliana ent*-copalyl diphosphate synthase proteins, were imported into the online software SWISSmodel by ExPASy (Waterhouse *et al.* 2018). SWISSmodel was used to do the initial predictive modelling of the protein structures. The template used for the modelling was the *Arabidopsis thaliana ent*-copalyl diphosphate synthase (PDB: 3PYA), this being the most identical protein sequence to that of the maize ent-copalyl diphosphate synthases, which the crystalline structure has been resolved. Final protein modelling and ligand docking was done using Maestro Version 12.5.139 (Schrödinger, Inc.). The protein structures were imported into Maestro and prepared using the protein preparation wizard. Pre-processing of the structures were done by the addition of hydrogen and disulphide bonds. Hydrogen bond assignments were optimised and protonation states were predicted using PROPKA (Olsson *et al.* 2011) at a pH of 7.0. The pH of a plant cell varies depending on the organelle in question, the cytosolic pH is typically 7.2 to 7.4, while the apoplast and vacuole exhibit a more acidic pH between 5 and 7. The pH of 7 was decided upon because the pH of the chloroplast's stroma, in which the ZmAn2 enzyme is found, is a pH of 7 to 8 (Ishizawa 2014, Su and Lai 2017). Finally, the enzymes were restrained and subjected to energy minimization using the OPLS3e force field (Harder *et al.* 2016). The ramachandram reports of these enzymes are displayed in Supplementary Figure 3.2. The ligand, geranly-geranyl diphosphate (universal GA precursor molecule), was prepared and minimised using the LigPrep function. Possible Het states were generated using Epik at a pH

of 7.0 +/- 2.0 and the OPLS3e force field was used for energy minimisation. Two deprotonated states of the ligand were generated (Supplementary Figure 3.1), the most deprotonated version was used for ligand docking as this version was the only state of the ligand which was predicted to be present at pH 7 to 8 (chloroplast's stromal pH in which reaction occurs). Ligand preparation data is summarised in Supplementary Table 3.2.

The catalytic sites on the enzymes were predicted using the SiteMap site analysis function using the default settings (Halgren 2007, 2009). Predicted binding sites with the highest site scores and which corresponded to amino acid sequence positions for the predicted catalytic sites were focused on. Receptor grid files were generated for each of the enzymes' sites using the receptor grid generation tool at a 20Å region around the centre of the identified pockets (Friesner *et al.* 2004, Halgren *et al.* 2004). A flexible docking was performed with the specific binding sites selected using Glide software (Friesner *et al.* 2004, Halgren *et al.* 2004). Docking of the most deprotonated state of geranyl-geranyl diphosphate (Supplementary Figure 3.1 B) in the binding sites of selected maize An2 and *Arabidopsis ent*-CPS enzymes were performed using the extra-precision (XP) function (Friesner *et al.* 2006). The binding affinity of the ligand to the binding sites were determined primarily by comparing the XP GScores generated from the ligand docking and were further investigated by observing the 2D ligand-protein interaction diagrams.

3.3 Results

3.3.1 Phylogenetic analysis of ent-copalyl diphosphate synthase

The phylogenetic analysis of the *ent*-copalyl diphosphate synthase amino acids sequences included sequences from 37 different *Z. mays* inbred lines, *P. virgatum*, *S. italica, T. aestivum*, *O. sativa*, *A. lyrata* and *A. thaliana* (Figure 3.2). Two *ent*-copalyl diphosphate synthase sequences (*ent*-CPS1 and *ent*-CPS2) were identified from *O. sativa* and included. Only singular *ent*-CPS amino acid sequences could be found for the other non-maize species listed. The An1 maize sequences were 56 – 60% identical to the An2 maize sequences. The An2 maize sequences across the different maize inbred lines were 94 – 100% identical.

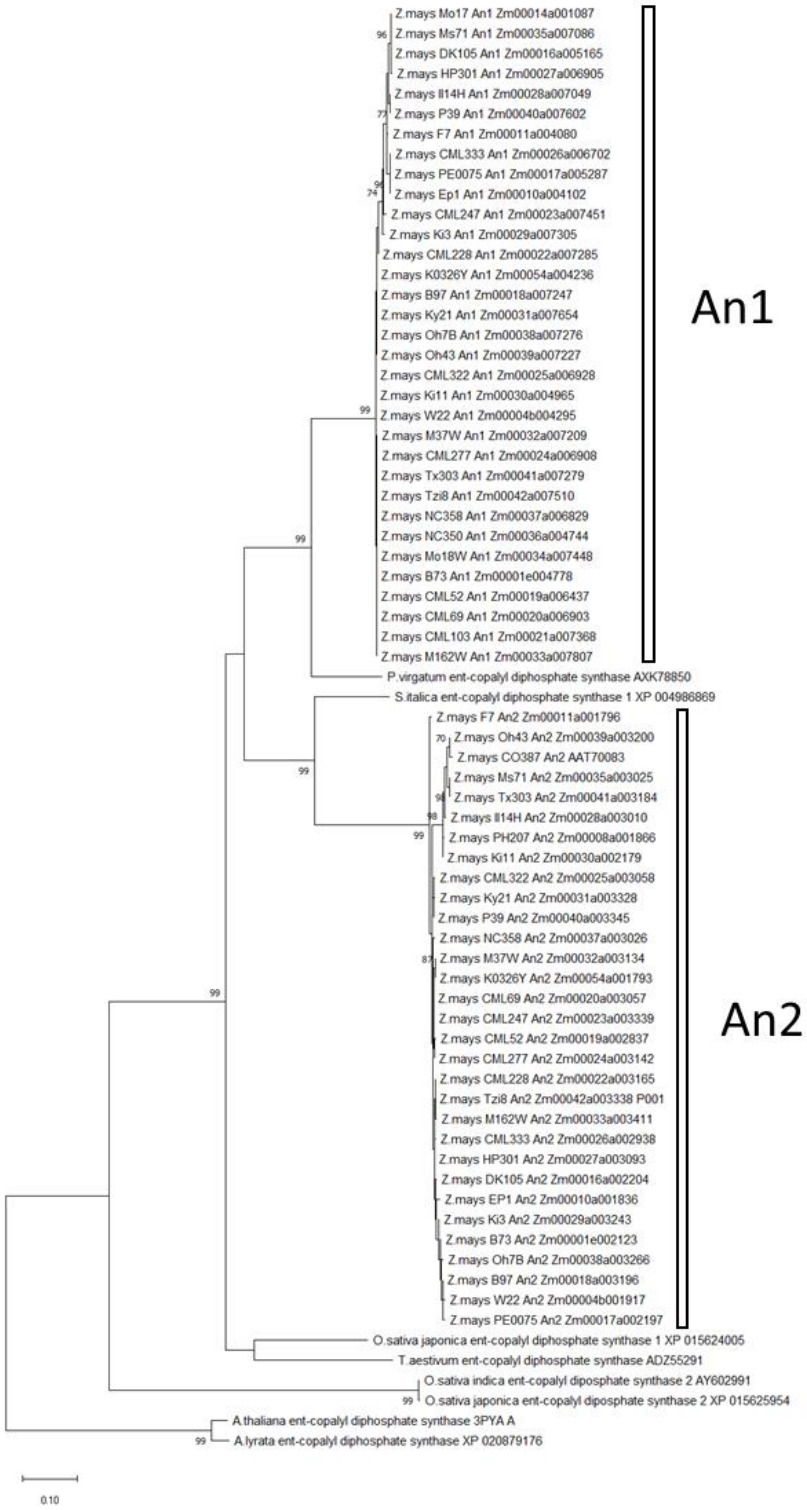


Figure 3.2: Maximum-likelihood phylogenetic analysis of *ent*-copalyl diphosphate synthase amino acid sequences from multiple maize (*Zea maydis*) genomes and other genera part of the Poaceae and Brassicaceae families, namely *Arabidopsis, Oryza, Triticum* and *Panicum*. The phylogenetic tree was generated using MEGAX, using the JTT model with 1000 boots trap iterations. Scale is 0.10 amino acid changes per 100 amino acids.

The *Z. mays* An1 and An2 proteins form two distinctive clades in the phylogenetic tree, An1 forming the top clade and An2 the bottom clade. Both clades are strongly supported with 99% bootstrap support. There is low bootstrap support within the clades to confidently group certain *Z. mays* An sequences together as being more or less similar. The two *ent*-CPS sequences belonging to the family Brassicaceae, namely *A. thaliana* and *A. lyrata*, form the outgroup and are used to root the tree. Although the maize An1 and An2 sequences are similar, the distinctive clades allow for confidence that none of the An2 sequences obtained are actually An1 proteins.

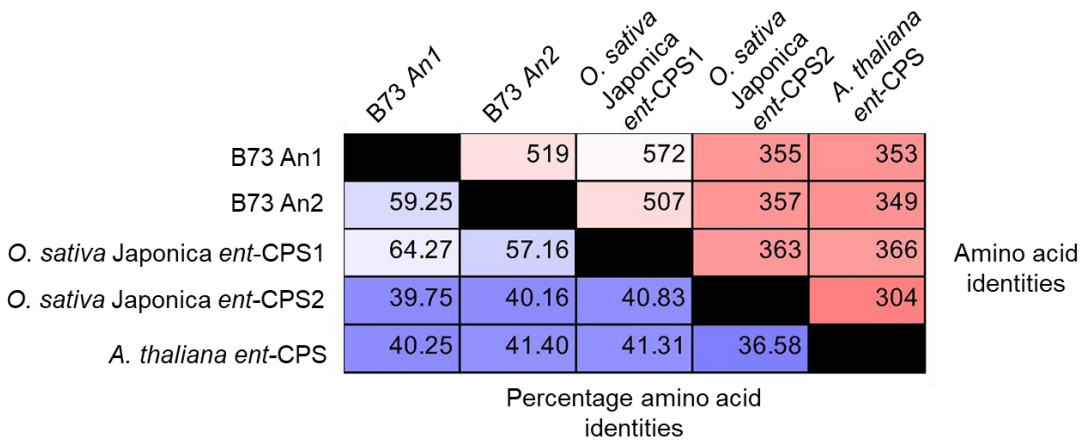


Figure 3.3: Pairwise comparison of *ent*-copalyl diphosphate sequences from *Z. mays* B73 inbred line, *O. sativa* Japonica and *A. thaliana*. The upper right comparison is the number of amino acid identities (red) and the bottom left comparison is the percentage identity of amino acids between the different sequences (blue).

**Note: The darker colours indicates lower amino acid identities and percentage identities.*

The pairwise comparison highlights the variation in amino acid sequences across the different enzymes and plant genera (Figure 3.3). The upper right comparison, illustrates the number of identical amino acids occurring at overlapping alignment positions between the sequences being compared. The bottom left comparison indicates the percentage of the identities illustrated in the upper comparison between the sequences being compared. The B73 An1 and An2 proteins are 59% identical in amino acid sequence. The B73 An1 amino acid sequence is more identical to the *O. sativa ent*-CPS1 sequence (64%), than it is to the B73

An2 sequence (59%). The *A. thaliana ent*-CPS and *O. sativa ent*-CPS2 sequences were the least similar to the rest of the sequences, with these sequences being 39 – 41% identical to the B73 An1 and An2 sequences.

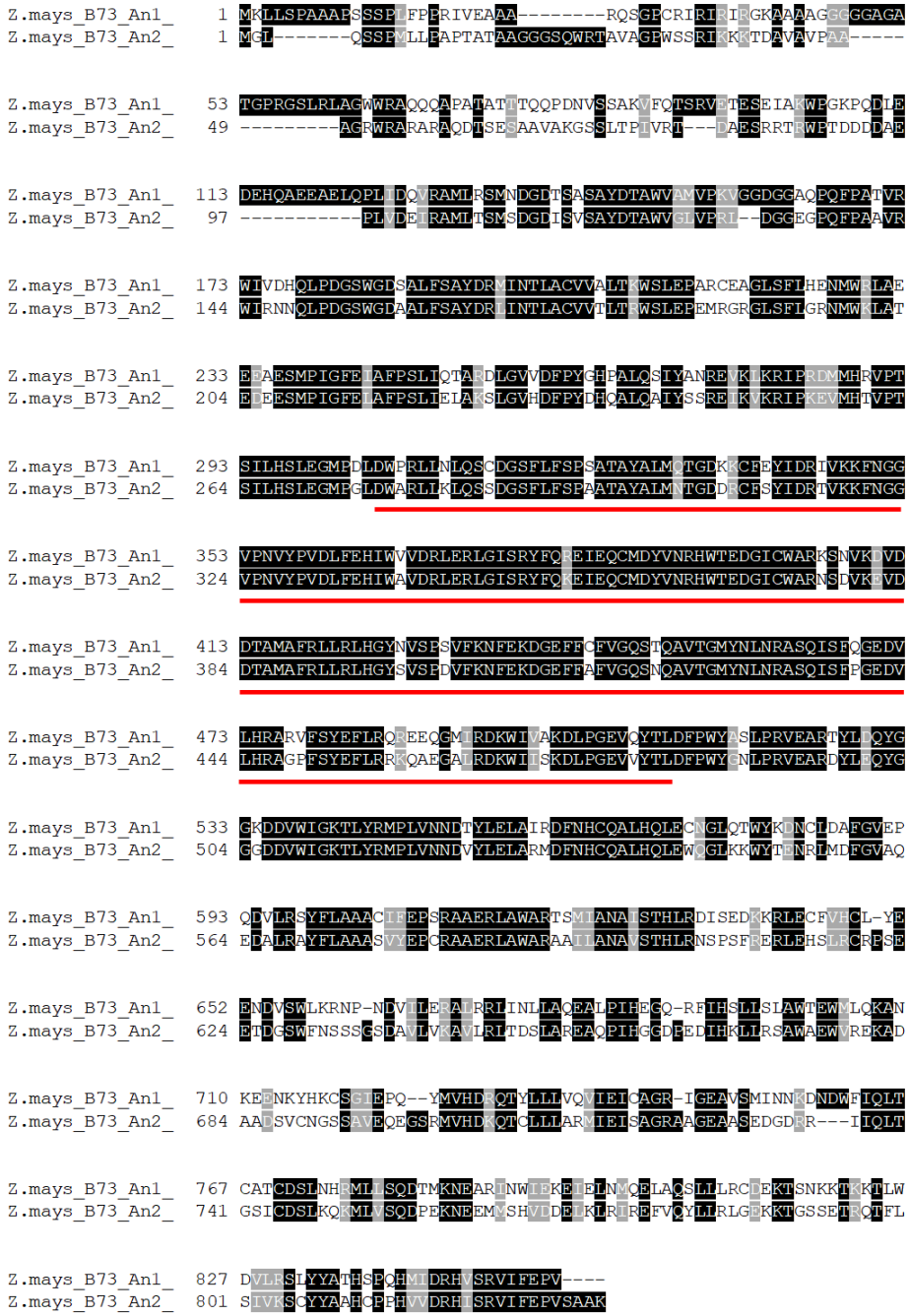


Figure 3.4: Amino acid sequence alignment of the *An1* and *An2* amino acid sequences of B73 Z. mays. The red line indicates the terpene synthase catalytic sites of both *ent*-copalyl diphosphate synthase proteins. Amino acid conservation was highlighted using boxshade. Black squares indicate conserved amino acids, light grey squares indicate conservative mutations and white squares indicate divergence. The dashes indicate gaps inserted in order to achieve sequence alignment. The numbering on the left side of the amino acid sequence corresponds to the number of amino acids.

The boxshade alignment (figure 3.4) illustrates the similarity of the An1 and An2 amino acid sequences of the model maize inbred line B73. The two protein sequences differ in length, the An1 protein being 827 aa long while the An2 protein is 801 aa long. The terpene synthase catalytic sites for both proteins, indicated as a red line, overlap and display high levels of conservation (black squares). The light grey and white squares illustrate conservative and divergent mutations, these are in the minority. The conservative mutations being mutations which result in similar amino acids, amino acids which have side chains of similar lengths and types (e.g. both having polar uncharged side chains). This alignment shows that although the enzymes separate into two distinct clades (Figure 3.2), there is high levels of conservation present in the catalytic sites of the An1 and An2 proteins of maize. This study further investigates solely the An2 protein of maize as it plays a critical role the synthesis of kauralexin maize defence compounds.

3.3.2 Catalytic site analysis of maize An2 proteins

Two major catalytic domains were identified when submitting the *Z. may* B73 An2 amino acid sequence to InterPro Scan, a terpene synthase and metal binding domain. The terpene synthase catalytic domain is positioned at amino acid number 268 and 497 along the 848 amino acid sequence (also indicated as the red line in Figure 3.4). The metal binding domain is positioned at amino acid number 541 and 697. The homologous superfamilies, which overlap the terpene synthase catalytic domain, are the terpene synthase superfamily and the terpenoid cyclase/protein prenyl transferase superfamily. The homologous superfamily which overlaps the metal binding domain is predominantly the isoprenoid synthase superfamily.

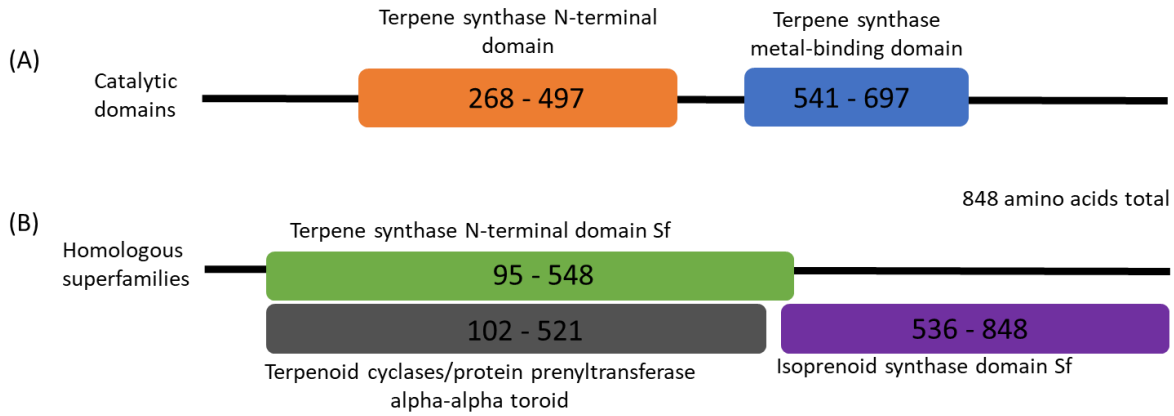


Figure 3.5: Catalytic domains (A) and homologous superfamilies (B) of the *Z. mays* B73 An2 amino acid sequence identified using InterPro Scan. The black line indicates the full amino acid sequences in a 5' to 3' orientation. In (A) two catalytic domains are identified, namely the terpene synthase domain (orange) and the terpene synthase metal-binding domain (blue). In (B) the homologous superfamilies are identified, namely the terpene synthase superfamily (green), the terpenoid cyclase/protein prenyl transferase superfamily (grey) and the isoprenoid synthase superfamily (purple).

Using the location of the terpene synthase catalytic domain (indicated as an orange segment in Figure 3.5 A), the terpene synthase catalytic domain amino acid sequence was used to perform the alignment and further phylogenetic analysis focusing on this catalytic sites (Figure 3.6).

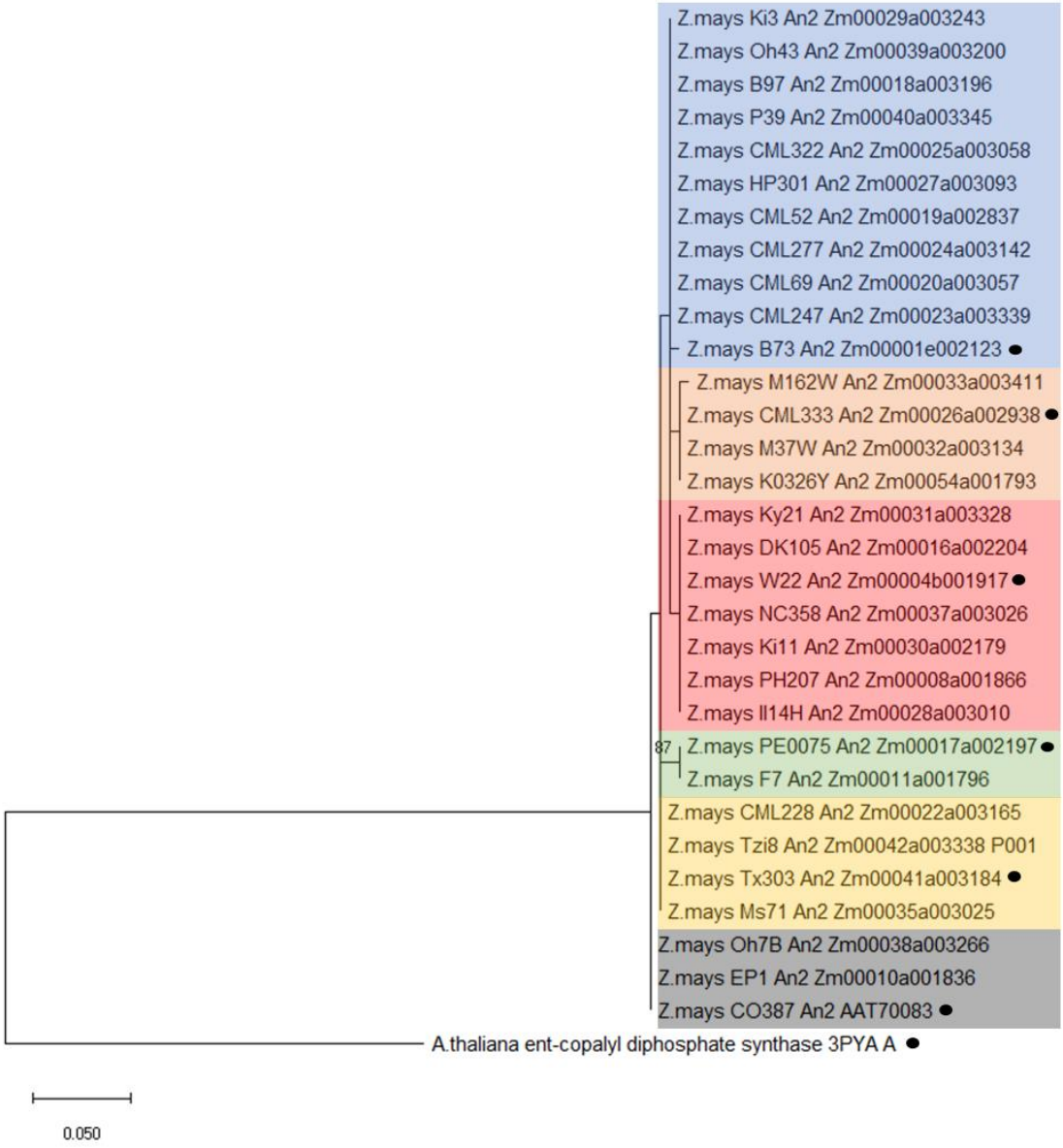


Figure 3.6: Maximum-likelihood phylogenetic analysis of the catalytic site from *ent*-copalyl diphosphate synthase 2 amino acid sequences from multiple maize genomes and the *A. thaliana ent*-CPS amino acid sequence as the outgroup. Lesser clades are highlighted in different colours. The black dots indicate amino acid sequences which were carried forward into the next analysis. The phylogenetic tree was generated using MEGAX, using the JTT model with 1000 bootstrap iterations. Scale is 0.05 amino acid changes per 100 amino acids.

A phylogenetic analysis using the terpene synthase catalytic sites of the *ent*-copalyl diphosphate synthase (An2) amino acid sequences was performed (Figure 3.6). The *Z. mays* An2 terpene synthase catalytic sites were grouped into 6 lesser clades, these groupings were highlighted using various colours for easy differentiation. The orange (51%) and green (87%)

were clades which exhibited the highest bootstrap support. All the other colour groups exhibited very low bootstrap support (< 40%). These colour groups (< 40% support) can be collapsed into a large group. These groupings were however kept in order to select individuals which could act as representatives for the variation present at the catalytic site across the multiple maize inbred lines. The differences between An2 sequences were investigated by choosing six *Z. mays* lines, one from each of the above coloured groups in order to include most of the variation between sequences and potential binding affinities of the ligand to the proteins. The amino acid sequences selected for further investigation were *Z. mays* An2 proteins for B73, W22, CO387, CML333, Tx303, PE0075 and the *A. thaliana ent*-CPS protein. The catalytic domains, namely the terpene synthase domains of the maize *An2* proteins were further investigated to determine how similar they were in amino acid sequence identity across the different maize lines. The six maize inbred line An2 proteins were aligned and the differences in amino acids were noted in the schematic below (Figure 3.7).

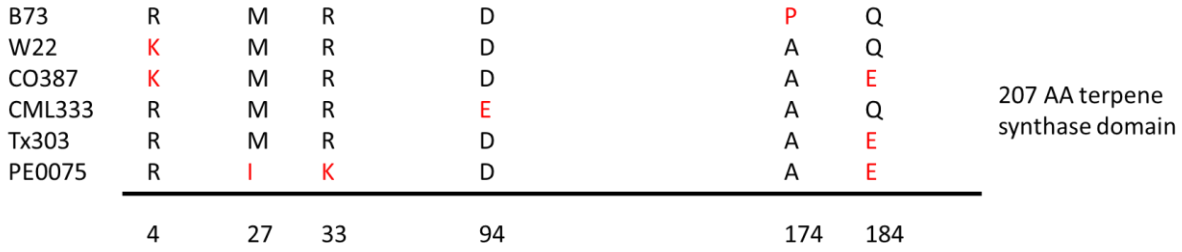


Figure 3.7: Schematic of the 207 amino acid terpene synthase domain of the *ent*-copalyl diphosphate synthase 2 (An2) protein of the 6 selected maize inbred lines. Amino acid differences between maize lines are highlighted in red and amino acid positions noted below.

The 207 amino acid long terpene synthase domain was highly conserved, 201 of the amino acid positions were identical across maize lines while only six positions differed. There is a minimum of one amino acid difference between each of the ZmAn2 catalytic site sequences across the maize inbred line. The six amino acid positions, which differed, is illustrated above in Figure 3.7.

The whole amino acid sequences of these 7 proteins were aligned and subject to pairwise comparison in order to determine the amino acid identities and percentage identities between sequences (Figure 3.8).

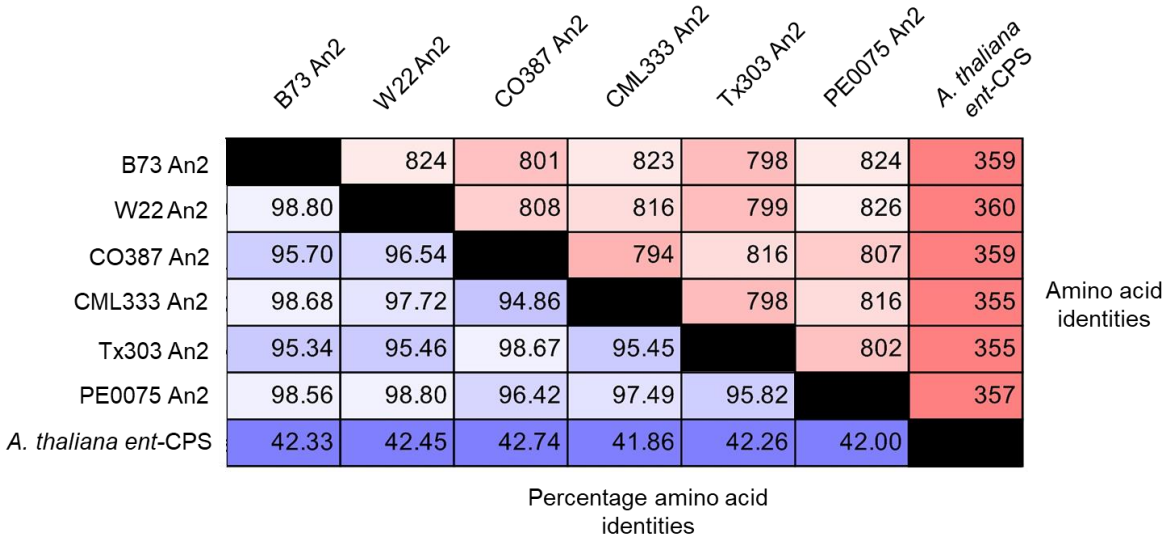


Figure 3.8: Pairwise comparison of *ent*-copalyl diphosphate synthase amino acid sequences of the six selected maize inbred lines and the *A. thaliana* outgroup. The upper right comparison is amino acid identities (red) and bottom left comparison is the percentage amino acid identities (blue).

Note: The darker colours indicates lower amino acid identities and percentage identities.

The percentage identities between the different maize inbred line An2 proteins showed high conservation (Figure 3.8). The highest percentage identity between maize inbred lines was that of the An2 proteins of B73 and W22 (98.8%); and W22 and PE0075 (98.8%). The lowest percentage identity between maize inbred line was that of the An2 proteins of CO387 and CML333 (94.86%) (Figure 3.8). The *Arabidopsis ent*-CPS protein was vastly different from that of the maize An1 and An2 proteins, with all maize proteins only being 42% identical in amino acid sequence identity to the *Arabidopsis ent*-CPS protein.

3.3.3 Protein modelling and ligand docking

The selected amino acid sequences were imported into Maestro. These enzymes were prepared using the protein preparation wizard. Examples of the 3D structures of these

enzymes can be seen below in Figure 3.9. Although the *A. thaliana ent*-CPS protein served as the template for the maize An2 proteins, notable differences are still present (for example the different α-helix formation in blue). The catalytic site region (yellow, orange and red) across the three enzymes appear to be highly similar.

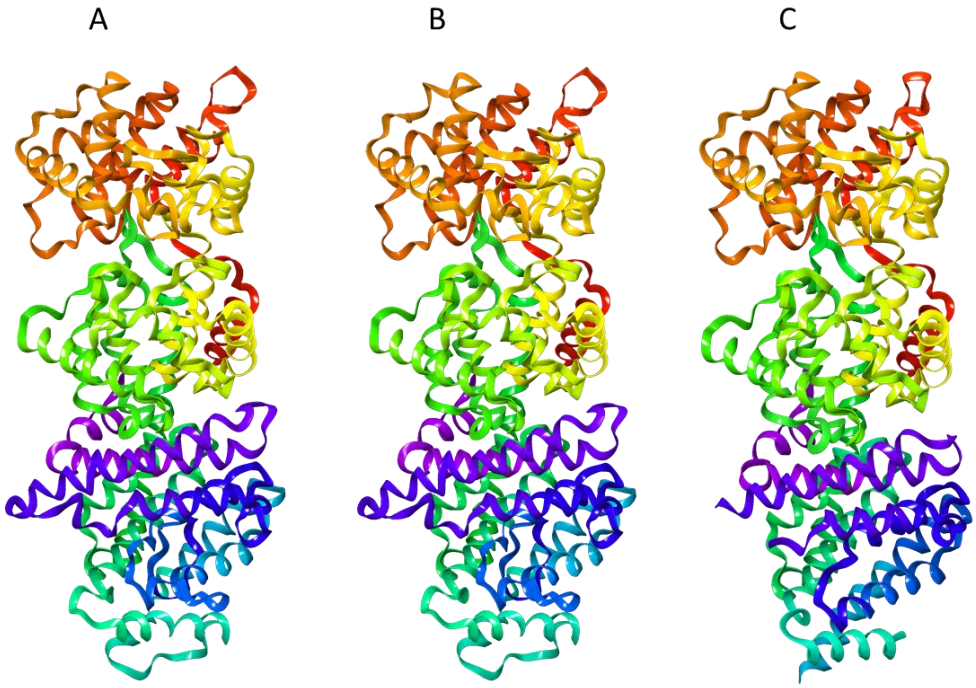


Figure 3.9: The ent-copalyl diphosphate structures (*An2*) of *Zea mays* inbred line B73 (A), W22 (B) and the *Arabidopsis thaliana ent*-copalyl diphosphate (C) which served as the template for protein modelling. The different colour of the ribbons indicate different secondary structures. These are the protein structures after the process of protein preparation.

A total of five binding sites were predicted on different parts of the protein structures. The volume and surface area details for these sites are displayed in Supplementary Table 3.3. The predicted sites with the highest scores (Supplementary table 3.3) and which corresponded to the amino acid sequence positions of the terpene catalytic sites were used for ligand docking. The GGPP ligand was docked successfully onto the selected ZmAn2 and *A. thaliana ent*-CPS proteins, the ligand for B73 ZmAn2 is illustrated below in Figure 3.10. The ligand docking results are summarised in Table 3.3.

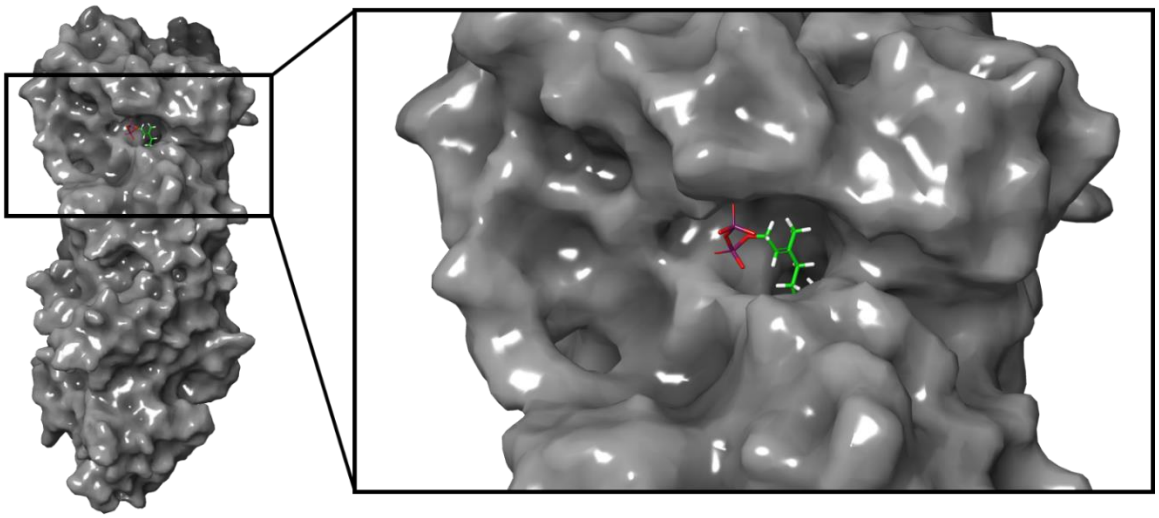


Figure 3.10: The docking of the ligand, GGPP, onto site 1 of the *Zea mays* B73 inbred line, An2 protein.

Table 3.1: The molecular docking results of geranyl-geranyl diphosphate within predicted binding site 1 of the maize and *Arabidopsis ent*-copalyl diphosphate synthase enzymes.

Genotype	XP GScore[a]	glide evdw[b]	glide ecoul[c]	glide energy[d]	glide einternal[e]	glide emodel[f]	XP HBond[g]
Z. mays B73	-5.281	-33.974	-16.109	-50.083	6.947	-67.264	-2.000
Z. mays W22	-5.722	-35.802	-13.291	-49.093	12.287	-66.062	-1.486
Z. mays CO387	-5.457	-35.013	-18.685	-53.698	8.547	-70.965	-1.840
Z. mays CML333	-5.681	-36.600	-14.264	-50.865	12.311	-67.680	-1.834
Z. mays Tx303	-5.051	-30.508	0.029	-30.479	11.750	-36.236	-2.163
Z. mays PE0075	-5.295	-36.709	-17.317	-54.026	12.060	-70.664	-1.523
A. thaliana	-10.735	-34.508	-22.758	-57.267	11.604	-73.743	-1.640

[a]*XP GScore: empirical scoring method for binding affinity (kcal/mol)*
[b]*Glide evdw: Van der Waals energy*
[c]*Glide ecoul: Coulomb energy*
[d]*Glide energy: Modified Coulomb-van der Waals interaction energy*
[e]*Glide einternal: Internal torsional energy*
[f]*Glide emodel: Model energy, Emodel*
[g]*XP HBond: Hydrogen bonding term in the GlideScore*

Among the ligand-receptor dockings, the enzyme which displayed the lowest XP Gscore for site 1 was *A. thaliana* (-10.735), the template enzyme for which the structure of the protein has been experimentally determined. The maize enzymes displayed little variation, the lowest XP GScore was *Z. mays* W22 with a score of -5.722. The highest XP GScore was that of *Z. mays* Tx303 with a score of -5.051.

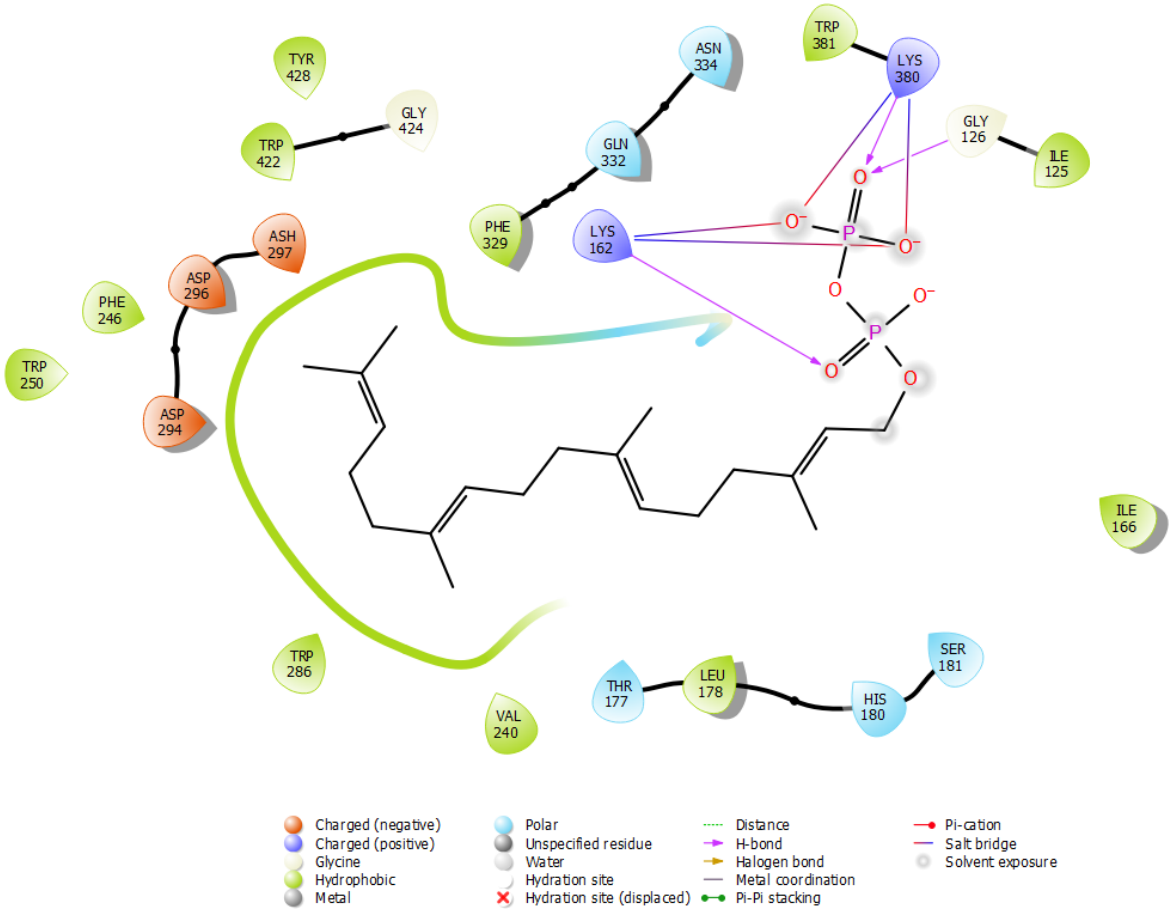


Figure 3.11: Two dimensional ligand-protein interaction diagram of the geranyl-geranyl diphosphate ligand in site 1 of the *A. thaliana ent*-CPS protein (XP GScore: -10.735).

In the ligand-protein interaction diagram of figure 3.11, the GGPP ligand is seen bound to the *A. thaliana ent*-CPS binding site. The amino acids of the protein which interact with the ligand predominantly are Lys380, Lys162 and Gly126. The amino acid Lys380 forms two salt bridges, 4.02 Å and 4.72 Å from the ligand, and a hydrogen bond 1.82 Å from the ligand. The amino acid Lys252 forms two salt bridges, 2.71 Å and 4.55 Å from the ligand, and a hydrogen bond 1.62 Å from the ligand. The amino acid Gly126 forms a hydrogen bond 2.57 Å from the ligand. The ligand-protein interaction illustrated in figure 3.12 and their respective summaries in table 2.4 are that of the specific ZmAn2 proteins tested from the various maize inbred lines.

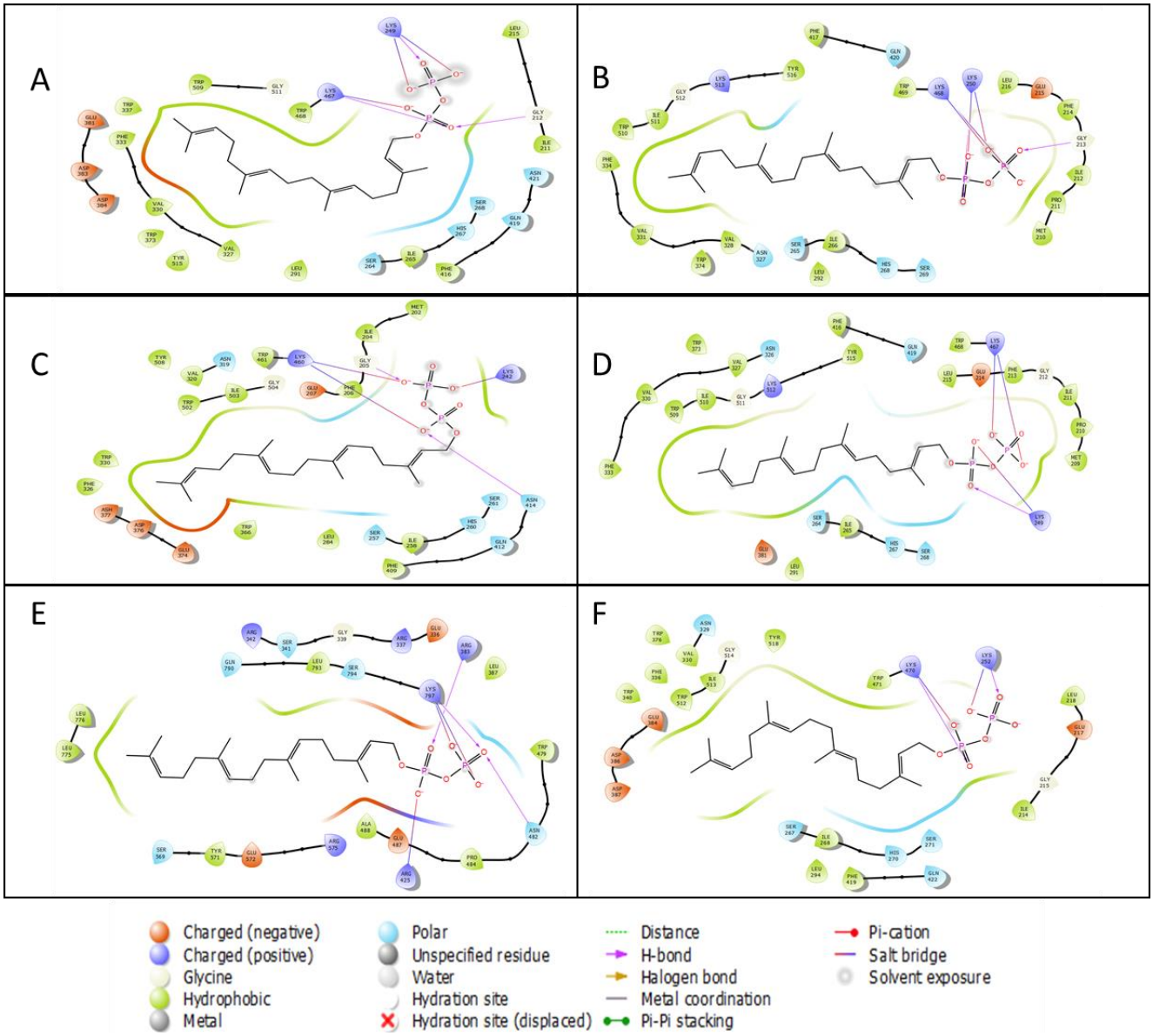


Figure 3.12: Two dimensional ligand-protein interaction diagrams of the geranyl-geranyl diphosphate ligand in site 1 of the ent-copalyl diphosphate synthase enzymes (An2) of the 6 *Z. mays* inbred lines. A - B73, B – W22, C – CO387, D – CML333, E – Tx303, F – PE0075.

Table 3.2: Summary of amino acid interactions with the geranyl-geranyl diphosphate ligand of *A. thaliana* and *Z. mays.*

	AA1	Bonds	Length (Å)	AA2	Bonds	Length (Å)	AA3	Bonds	Length (Å)	AA4	Bonds	Length (Å)
A thaliana	Lys162	salt bridge	2,71	Lys380	salt bridge	4,02	Gly162	H bond	2,57			
		salt bridge	4,55		salt bridge	4,72						
		H bond	1,62		H bond	1,82						
B73	Lys249	salt bridge	2,96	Lys467	salt bridge	3,87	Gly212	H bond	2,42			
		salt bridge	4,77		H bond	1,64						
		H bond	1,81									
W22	Lys250	salt bridge	4,65	Lys468	salt bridge	2,96	Gly213	H bond	1,95			
		salt bridge	3,68		salt bridge	2,95						
		H bond	1,8									
CO387	Lys242	salt bridge	2,78	Lys460	salt bridge	3,48	Gly205	H bond	1,77	Asn414	H bond	1,94
					salt bridge	4,68						
CML333	Lys249	salt bridge	4,3	Lys467	salt bridge	2,66						
		H bond	1,74		salt bridge	3,23						
Tx303	Lys797	salt bridge	3,49	Arg383	H bond	2,33	Arg425	salt bridge	3,07	Asn482	H bond	1,88
		salt bridge	4,64									
		H bond	2,11									
PE0075	Lys252	salt bridge	4,06	Lys470	salt bridge	4,08						
		H bond	1,62		H bond	1,68						

The maize inbred lines B73, W22 and CO387 conform to the *A. thaliana ent*-CPS protein interaction diagram. These maize lines all have two lysine and one glycine amino acids which interacts with the GGPP ligand. The maize line CO387 has an additional asparagine amino acid which interacts with the ligand forming an H bond (Figure 3.12 C). The maize inbred lines CML333 and PE0075 both have two lysine's, similar to that of *A. thaliana*, which interact with the ligand (Figure 3.12 D & F). The maize line Tx303 has the most unique array of amino acids which interact with the ligand (Figure 3.12 E). It is the only protein to have only one lysine interacting with the ligand, it has an asparagine amino acid similar to that of CO387, and two arginine amino acids which uniquely form interactions with the ligand.

3.4 Discussion

The kauralexin and gibberellic acid synthesis pathways are closely related, in that the enzymes which make up the kauralexin synthesis pathway are products of gene duplication events within the GA synthesis pathway (Ding *et al.* 2019). These pathways utilise the same terpenoid precursor molecule GGPP and yet synthesise compounds, which perform very different functions, namely plant development (GA) and plant defence (kauralexins) (Harris *et al.* 2005, Schmelz *et al.* 2011, Vaughan *et al.* 2015, Christensen *et al.* 2018). The two *ent*-copalyl diphosphate synthase enzymes, ZmAn1 and ZmAn2, are involved in the first steps of the GA and kauralexin synthesis, respectively, catalysing the cyclisation of GGPP to CDP (Ding *et al.* 2019). The *ent*-copalyl diphosphate synthase enzyme is present in most plants but gene duplications of this enzyme resulting in new secondary metabolite pathways (i.e. An2 / *ent*-CPS2) are only present in few specific species such as *Z. mays* and *O. sativa* (Prisic *et al.* 2004, Schmelz *et al.* 2011, Toyomasu *et al.* 2015).

Further investigation in this study of *ent*-copalyl diphosphate synthase enzymes across multiple maize inbred lines and other species within the family Poaceae and Brassicaceae was performed. Phylogenetic analysis of these proteins show two separate clades, both with high bootstrap support (99%), for the ZmAn1 and ZmAn2 proteins of the 37 maize inbred lines. The two separate clades indicated clear and reoccurring differences between these two proteins at an amino acid sequence level. The *ent*-copalyl diphosphate synthase (*ent*-CPS) proteins of the non-*Zea* sequences did not cluster together indicating variation in *ent*-CPS proteins across the different genera tested. Greater percentage identity between ZmAn1 and *O. sativa ent*-CPS1 sequences (64%) compared to ZmAn2 and *O. sativa ent*-CPS1 sequences (57%) was observed in the pairwise comparison. This is most likely due to the ZmAn1 and *O. sativa ent*-CPS1 proteins both forming part of GA synthesis pathways and these genera being part of the same order (Poales) and family (Poaceae) (Sakamoto *et al.* 2004, Nagata *et al.*

2016, Lemke *et al.* 2019). Low percentage identity was however observed between ZmAn2 and *O. sativa ent*-CPS2 proteins (39%). Although both enzymes *ent*-copalyl diphosphate synthases which are part of secondary metabolite pathways – each results in compounds which are significantly different which may account for the lack of similarity between the proteins. In maize, this results in kauralexins and dolabralexins, and in rice this forms oryzalexins. Direct comparison of model maize inbred line B73 ZmAn1 and ZmAn2 protein sequences was performed by means of alignment. The alignment shows how the terpene synthase catalytic sites are highly similar, this is likely due to the identical chemical reaction they catalyse. The key difference between these enzymes, An1 and An2, are the synthesis pathways they form part of and that they are differentially regulated. The *An1* gene is expressed in new growth points where GA synthesis is required while *An2* is expressed in response to stressors such as pathogen invasion, insect herbivory or drought (Dafoe *et al.* 2011, Schmelz *et al.* 2011, Nelissen *et al.* 2012, Vaughan *et al.* 2015). These results illustrate that although the two maize proteins catalyse the same chemical reactions, they contain distinct amino acid differences when comparing the entire sequences which may be linked to differences noted above, but there are few amino acid differences when comparing the catalytic sites.

The amino acid sequences of the terpene synthase catalytic sites were extracted for all the maize inbred line An2 and *A. thaliana ent*-CPS (template protein for homology modelling) proteins. Phylogenetic analysis of these catalytic site sequences, using the *A. thaliana ent*-CPS catalytic site sequence as the outgroup, resulted in a clades forming which had low bootstrap support (<70%). The *A. thaliana ent*-CPS protein is utilised instead of other *ent*-CPS proteins from closer related Poaceae genera, as the *Arabidopsis* protein is the only one which has a crystalline structure and can serve as the template protein to construct the maize An2 homology models. The low support for the clades of ZmAn2 catalytic site sequences means that there are few differences between these sequences which are reoccurring across the maize lines that result in repeatable clade formation from multiple resampling iterations (1000

bootstrap iterations). Although the clades are not well supported, the purpose of this phylogenetic analysis was to group ZmAn2 sequences with similar catalytic site sequences, choosing representatives from each clade in order to maximise the variation of catalytic sites tested in later protein docking studies. From the 6 major clades of the phylogenetic tree, the representative maize lines chosen from each of these clades were B73, W22, CO387, CML333, Tx303, PE0075. The reasons for choosing these inbred lines is that B73 is a model maize inbred line and its ZmAn2 protein sequence was submitted to perform the BLAST query to identify the other ZmAn2 protein sequences; W22, as it is the maize line used in this study's fall armyworm assays; CO387, as its sequence is confirmed with sequencing data from a published article (Harris *et al.* 2005); and the remaining inbred lines, CML333, Tx303 and PE0075, were chosen at random from the clades. The percentage identities between whole amino acid sequences of the selected ZmAn2 proteins are above 94%, indicating high conservation of this enzyme (Figure 3.8). Further investigation into the terpene synthase catalytic sites of these enzymes through alignment, revealed a minimum of one amino acid difference was present between each of the catalytic sites. One amino acid change in proteins could result in significant structural changes and function, especially if the change is found in the catalytic sites (Schaefer and Rost 2012).

The 3D protein structures of the selected maize lines and *Arabidopsis* template protein were investigated to determine whether the amino acid changes observed could result in structural changes which have the potential to alter total kauralexin synthesis levels. Homology modelling of the ZmAn2 proteins was performed using the *Arabidopsis ent*-CPS protein, the closest available protein structure as the template. Although the template is not high in sequence identity (~42%), homology models of medium accuracy are built from protein templates which display 30 – 50% sequence identity (Tastan Bishop *et al.* 2008). The protein structures and ligand protonation states were optimised for pH7. This pH was selected to mimic that of the neutral to basic pH found in the chloroplast's stroma where the enzyme is found (Bishop *et al.* 1972, Ishizawa 2014, Su and Lai 2017). Of the catalytic sites generated

by SiteMap, site 1 was the only site across the 7 *ent*-copalyl diphosphate synthase proteins in which amino acid composition aligned with that of the catalytic site sequences determined earlier in the study. This is also confirmed by site 1 consistently displaying the highest site scores compared to other predicted sites.

The ligand docking (using GGPP as the ligand) on the homology models and *Arabidopsis ent*-CPS protein was performed in order to determine whether there was any differences in binding affinity between the protein models. The protein structure to which the ligand displayed the highest binding affinity (most negative XP GScore), was that of the *Arabidopsis ent*-CPS protein. This was expected as this is the only protein structure for which the crystalline structure has been determined (Köksal *et al.* 2011). The maize proteins to which the ligand displayed the highest binding affinity to was the W22 and CML333 An2 proteins. This result linked well with the studies that determined total kauralexin levels across multiple maize inbred lines where W22 and CML333 were among the highest kauralexin synthesising lines (Schmelz *et al.* 2011, Ding *et al.* 2019). The ligand displayed lower binding affinities to the CO387, PE0075, B73 and Tx303 An2 proteins (Tx303 being the lowest). The maize lines B73 and Tx303 both also linked well with the studies containing total kauralexin level data, with B73 and Tx303 forming the middle to lower end of kauralexin synthesizing lines. The ligand-protein interaction diagrams of the maize lines W22, CML333, CO387, PE0075 and B73 generally conform to that of the template *Arabidopsis ent*-CPS protein, in that most of these proteins have two lysines and one glycine which interacts with the ligand. The outlier ZmAn2 protein, Tx303, is the only protein which does not follow this pattern, the reason for this is unclear but may be due to errors in either the sequence data or in the final protein structure. The amino acid glutamine at position 184 on the An2 catalytic site sequences (position 460 on the whole An2 protein sequences) is present in W22, CML333 and B73. The maize inbred lines CO387, Tx303 and PE0075 exhibit a glutamic acid at this position, this difference may be part of the reason for the differences in total kauralexins observed between these maize lines. The linking of this study's binding affinity data to published experimental results is promising and highlights

the influence ZmAn2 proteins may have over the total kauralexin synthesis potential of maize plants.

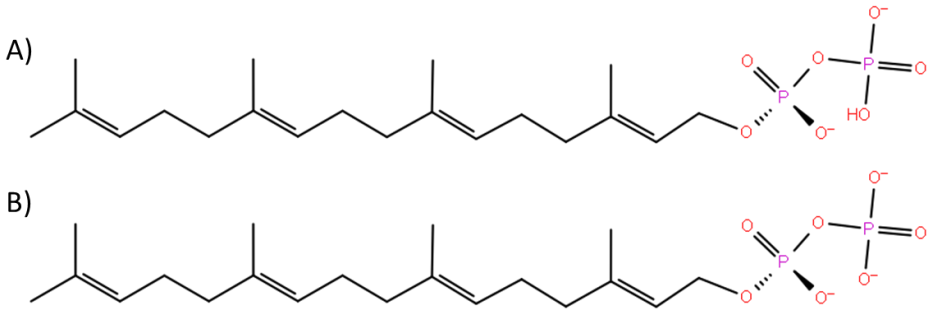
3.5 Supplementary information

Supplementary Table 3.1: The accession numbers and sources of the e*nt*-copalyl diphosphate synthase amino acid sequences used the phylogenetic analysis (Figure 2.2).

Plant Species	Protein	Protein version	Accession Number	Online Source
Z. mays B73	An1	P001	Zm00001e004778	maizeGDB
Z. mays B97	An1	P001	Zm00018a00724	maizeGDB
Z. mays CML103	An1	P001	Zm00021a007368	maizeGDB
Z. mays CML228	An1	P002	Zm00022a007285	maizeGDB
Z. mays CML247	An1	P001	Zm00023a007451	maizeGDB
Z. mays CML277	An1	P001	Zm00024a006908	maizeGDB
Z. mays CML322	An1	P003	Zm00025a006928	maizeGDB
Z. mays CML333	An1	P001	Zm00026a006702	maizeGDB
Z. mays CML52	An1	P001	Zm00019a006437	maizeGDB
Z. mays CML69	An1	P001	Zm00020a006903	maizeGDB
Z. mays HP301	An1	P005	Zm00027a006905	maizeGDB
Z. mays Il14H	An1	P001	Zm00028a007049	maizeGDB
Z. mays Ki11	An1	P002	Zm00030a004965	maizeGDB
Z. mays Ki3	An1	P003	Zm00029a007305	maizeGDB
Z. mays Ky21	An1	P002	Zm00031a007654	maizeGDB
Z. mays M162W	An1	P001	Zm00033a007807	maizeGDB
Z. mays M37W	An1	P001	Zm00032a007209	maizeGDB
Z. mays Mo18W	An1	P001	Zm00034a007448	maizeGDB
Z. mays Ms71	An1	P002	Zm00035a007086	maizeGDB
Z. mays NC350	An1	P001	Zm00036a004744	maizeGDB
Z. mays NC358	An1	P001	Zm00037a006829	maizeGDB
Z. mays Oh43	An1	P002	Zm00039a007227	maizeGDB
Z. mays Oh7B	An1	P001	Zm00038a007276	maizeGDB
Z. mays P39	An1	P003	Zm00040a007602	maizeGDB
Z. mays Tx303	An1	P002	Zm00041a007279	maizeGDB
Z. mays Tzi8	An1	P001	Zm00042a007510	maizeGDB
Z. mays EP1	An1	P001	Zm00010a004102	maizeGDB
Z. mays F7	An1	P001	Zm00011a004080	maizeGDB
Z. mays DK105	An1	P001	Zm00016a005165	maizeGDB
Z. mays K0326Y	An1	P001	Zm00054a004236	maizeGDB
Z. mays Mo17	An1	P001	Zm00014a001087	maizeGDB
Z. mays PE0075	An1	P001	Zm00017a005287	maizeGDB
Z. mays PH207	An1	P001	Zm00008a004375	maizeGDB
Z. mays W22	An1	P002	Zm00004b004295	maizeGDB
Z. mays B73	An2	P001	Zm00001e002123	maizeGDB
Z. mays B97	An2	P001	Zm00018a003196	maizeGDB
Z. mays CML228	An2	P001	Zm00022a003165	maizeGDB
Z. mays CML247	An2	P001	Zm00023a003339	maizeGDB
Z. mays CML277	An2	P001	Zm00024a003142	maizeGDB

Z. mays CML322	An2	P001	Zm00025a003058	maizeGDB
Z. mays CML333	An2	P001	Zm00026a002938	maizeGDB
Z. mays CML52	An2	P001	Zm00019a002837	maizeGDB
Z. mays CML69	An2	P001	Zm00020a003057	maizeGDB
Z. mays HP301	An2	P001	Zm00027a003093	maizeGDB
Z. mays Il14H	An2	P001	Zm00028a003010	maizeGDB
Z. mays Ki11	An2	P001	Zm00030a002179	maizeGDB
Z. mays Ki3	An2	P001	Zm00029a003243	maizeGDB
Z. mays Ky21	An2	P002	Zm00031a003328	maizeGDB
Z. mays M162W	An2	P001	Zm00033a003411	maizeGDB
Z. mays M37W	An2	P001	Zm00032a003134	maizeGDB
Z. mays Ms71	An2	P003	Zm00035a003025	maizeGDB
Z. mays NC358	An2	P001	Zm00037a003026	maizeGDB
Z. mays Oh43	An2	P001	Zm00039a003200	maizeGDB
Z. mays Oh7B	An2	P001	Zm00038a003266	maizeGDB
Z. mays P39	An2	P001	Zm00040a003345	maizeGDB
Z. mays Tx303	An2	P002	Zm00041a003184	maizeGDB
Z. mays Tzi8	An2	P001	Zm00042a003338	maizeGDB
Z. mays DK105	An2	P001	Zm00016a002204	maizeGDB
Z. mays K0326Y	An2	P001	Zm00054a001793	maizeGDB
Z. mays Mo17	An2	P001	Zm00014a025659	maizeGDB
Z. mays EP1	An2	P001	Zm00010a001836	maizeGDB
Z. mays F7	An2	P001	Zm00011a001796	maizeGDB
Z. mays PE0075	An2	P001	Zm00017a002197	maizeGDB
Z. mays PH207	An2	P001	Zm00008a001866	maizeGDB
Z. mays W22	An2	P001	Zm00004b001917	maizeGDB
S. italica	*ent*-CPS1	XP_004986869.2	XP004986869	GenBank (NCBI)
O. sativa japonica	*ent*-CPS1	XP_015624005.1	XP015624005	GenBank (NCBI)
O. sativa japonica	*ent*-CPS2	XP_015625954.1	XP_015625954	GenBank (NCBI)
O. sativa indica	*ent*-CPS2	AY602991.1	AY602991	GenBank (NCBI)
P. virgatum	*ent*-CPS1	AXK78850.1	AXK78850	GenBank (NCBI)
T. aestivum	*ent*-CPS1	ADZ55291.1	ADZ55291	GenBank (NCBI)
A. lyrata	*ent*-CPS	XP_020879176.1	XP020879176	GenBank (NCBI)
A. thaliana	*ent*-CPS	3PYA_A	3PYA	GenBank (NCBI)

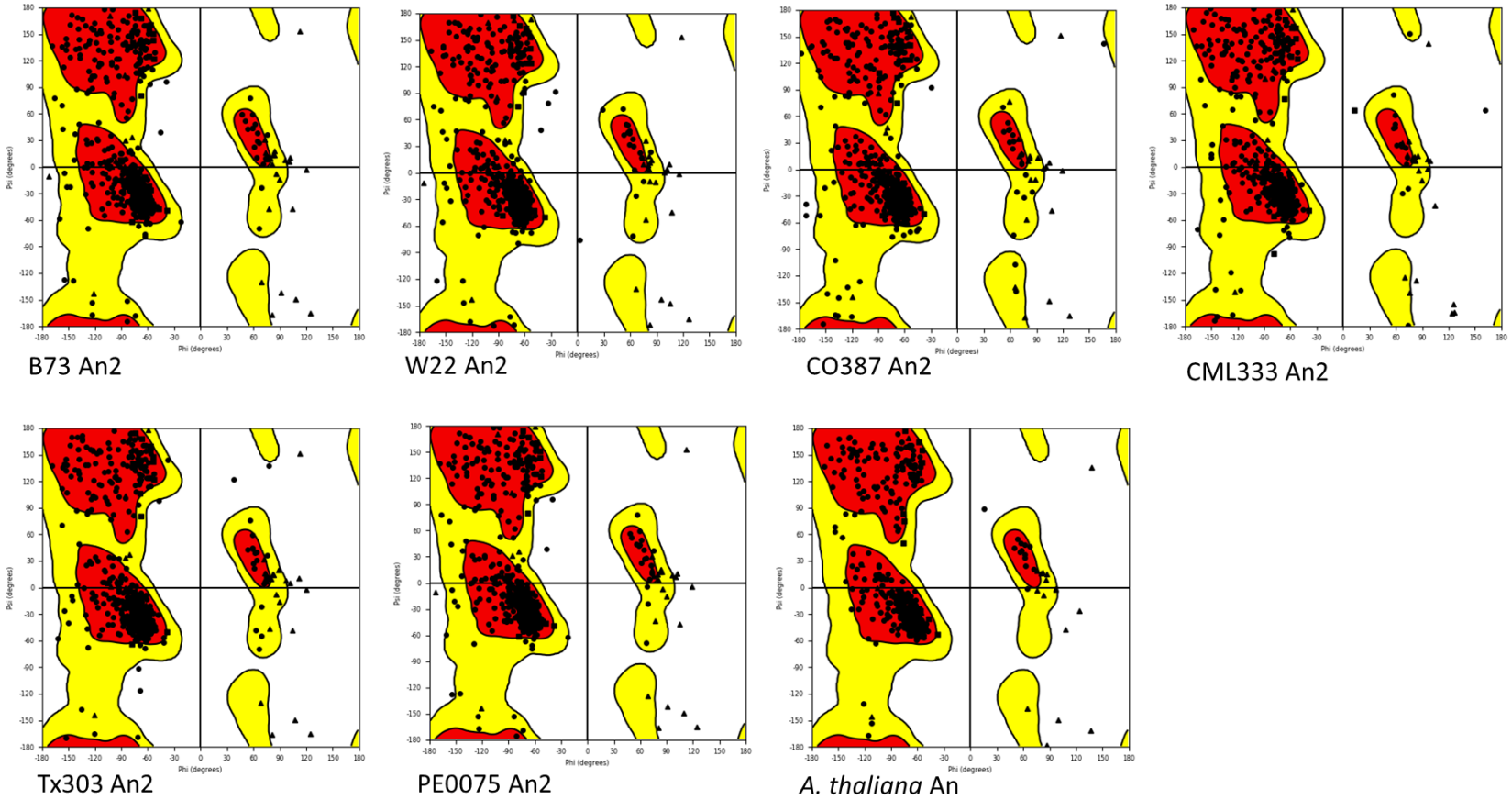
**Ent-copalyl diphosphate enzymes of non-Zea species which do not have a number associated are present as single copies of this enzyme in these genera.*



Supplementary Figure 3.1: The 2D structures of the two deprotonated forms of geranyl-geranyl diphosphate generated by the ligand preparation process.

Supplementary Table 3.2: Ligand preparation data of the ligand geranly-geranyl diphosphate.

Ligand name	Ligand version	mmshare version	Ionization Penalty	Ionization Penalty Charging	Ionization Penalty Neutral	State Penalty	Charging Adjusted Penalty	Tot Q	Tot abs Q	Energy
Geranyl-geranyl diphosphate	A	51139	1.1751	0.0003	1.1748	1.2382	1.238	-2	2	-22.033
	B	51139	0.0891	0.0891	0.0000	0.0782	0.167	-3	3	-11.325



Supplementary Figure 3.2: Ramachandram plots of maize An2 proteins and the *A. thaliana* CPS protein after the proteins were processed using the protein preparation wizard. Each black dot represents an amino acid and it's placement on the plot dependent on its phi (φ)and psi (ψ) angles (x and y-axis respectively). These plots use the torsional angles of the amino acids in order to predict the secondary structures present. The red areas on the plots highlight amino acids which form specific secondary structures, the red area to the centre left of each of the plots represents right-handed alpha helices, top left represents beta-sheets and the smaller centre right area represents left handed alpha helices.

Supplementary Table 3.3: SiteMap analysis of *Z. mays* An2 and *A. thaliana* CPS proteins.

Genotype	Site number	SiteScore[a]	Size[b]	Dscore[c]	Volume[d]	Exposure[e]	Enclosure[f]	Contact[g]	Phobic[h]	Philic[i]	balance[j]
Z. mays B73	1	373	1.042	960.400	0.539	0.743	0.945	0.485	1.054	0.460	0.834
	2	195	0.989	687.372	0.458	0.880	1.056	0.276	1.469	0.188	0.775
	3	164	1.020	332.367	0.376	0.879	1.137	0.466	1.371	0.340	0.555
	4	147	1.040	581.042	0.569	0.755	0.926	0.457	1.083	0.422	0.869
	5	104	0.967	271.656	0.559	0.728	0.987	0.335	1.257	0.267	0.708
Z. mays W22	1	417	1.033	1018.024	0.543	0.751	0.937	0.557	1.095	0.509	0.758
	2	252	1.002	855.099	0.517	0.800	0.999	0.348	1.283	0.271	0.848
	3	168	1.017	326.193	0.419	0.825	1.055	0.514	1.282	0.401	0.524
	4	133	1.069	499.751	0.543	0.777	1.018	0.778	1.036	0.751	0.748
	5	114	0.976	269.941	0.517	0.726	0.981	0.338	1.227	0.275	0.677
Z. mays CO387	1	424	1.001	1061.242	0.547	0.735	0.933	0.424	1.165	0.364	0.732
	2	294	1.054	957.656	0.557	0.737	0.953	0.700	1.007	0.695	1.084
	3	200	1.041	453.103	0.520	0.770	0.957	0.521	1.105	0.471	0.806
	4	186	0.985	646.212	0.488	0.881	1.092	0.196	1.482	0.132	0.729
	5	107	0.947	279.202	0.478	0.774	1.097	0.372	1.402	0.265	0.644
Z. mays CML333	1	343	1.029	797.475	0.471	0.794	0.997	0.612	1.190	0.514	0.826
	2	329	1.044	1181.635	0.533	0.782	1.015	0.544	1.122	0.485	0.740
	3	74	0.907	295.323	0.688	0.683	0.798	0.254	0.956	0.266	0.684
	4	69	0.832	234.269	0.737	0.620	0.734	0.125	0.984	0.127	0.786
	5	62	0.762	130.340	0.650	0.591	0.767	0.090	1.022	0.088	1.607
Z. mays Tx303	1	347	1.004	1054.382	0.439	0.824	1.071	0.409	1.323	0.309	0.796
	2	341	0.998	788.214	0.469	0.800	1.017	0.513	1.293	0.396	0.627
	3	224	1.140	404.397	0.406	0.818	1.095	1.489	0.894	1.665	1.519
	4	202	1.030	374.556	0.308	0.894	1.210	1.170	1.370	0.854	0.772
	5	129	1.035	340.256	0.554	0.730	0.987	0.715	1.050	0.681	1.219
Z. mays PE0075	1	425	1.019	1030.372	0.523	0.755	0.965	0.431	1.148	0.376	0.709
	2	182	1.014	682.570	0.481	0.880	1.063	0.295	1.391	0.212	0.822

	3	164	1.015	328.594	0.362	0.878	1.153	0.453	1.387	0.327	0.531
	4	147	1.043	579.327	0.561	0.755	0.926	0.483	1.073	0.450	0.887
	5	104	1.007	286.405	0.636	0.693	0.840	0.332	1.070	0.310	0.661
	1	191	1.165	536.109	0.454	0.851	1.025	1.436	0.875	1.641	0.597
	2	110	1.032	678.797	0.607	0.778	0.887	0.276	1.148	0.241	0.696
A. thaliana	3	89	1.035	226.380	0.440	0.837	1.131	0.878	1.084	0.810	1.751
	4	60	0.881	238.385	0.595	0.664	0.886	0.841	0.751	1.119	3.416
	5	39	0.695	148.176	0.519	0.717	1.032	0.115	0.987	0.116	0.755

A – Score determined generated using factors such as site points, enclosure and hydrophilicity.
B – Size of site measured in Angstrom (Å)
C – Druggability score
D – Volume of site measured in Angstrom (Å)
E – Measurement in Angstrom (Å) of how exposed the site is to solvent
F – Measurement in Angstrom (Å) of how enclosed the site is from solvent
G – Strength average site point interacts with surrounding receptor
H – Score defining hydrophobicity character of the site
I – Score defining hydrophilicity character of the site
J – Ratio of the hydrophobicity and hydrophilicity scores of the site

3.6 References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Bishop, D., K. S. Andersen, and R. M. Smillie. 1972. pH dependence and cofactor requirements of photochemical reactions in maize chloroplasts. Plant Physiology 50: 774-777.

Block, A. K., M. M. Vaughan, E. A. Schmelz, and S. A. Christensen. 2019. Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). Planta 249: 21-30.

Blum, M., H.-Y. Chang, S. Chuguransky, T. Grego, S. Kandasaamy, A. Mitchell*, et al.* 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Research 49: D344-D354.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17: 540-552.

Christensen, S. A., J. Sims, M. M. Vaughan, C. Hunter, A. Block, D. Willett*, et al.* 2018. Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize. Journal of Experimental Botany 69: 1693-1705.

Dafoe, N. J., A. Huffaker, M. M. Vaughan, A. J. Duehl, P. E. Teal, and E. A. Schmelz. 2011. Rapidly induced chemical defenses in maize stems and their effects on short-term growth of *Ostrinia nubilalis*. Journal of Chemical Ecology 37: 984-991.

Ding, Y., K. M. Murphy, E. Poretsky, S. Mafu, B. Yang, S. N. Char*, et al.* 2019. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. Nature Plants 5: 1043-1056.

Friesner, R. A., R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren*, et al.* 2006. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. Journal of Medicinal Chemistry 49: 6177-6196.

Friesner, R. A., J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz*, et al.* 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry 47: 1739-1749.

Gao, Y., R. B. Honzatko, and R. J. Peters. 2012. Terpenoid synthase structures: a so far incomplete view of complex catalysis. Natural Product Reports 29: 1153-1175.

Gupta, R., and S. Chakrabarty. 2013. Gibberellic acid in plant: still a mystery unresolved. Plant Signaling & Behavior 8: e25504.

Halgren, T. 2007. New method for fast and accurate binding-site identification and analysis. Chemical Biology & Drug Design 69: 146-148.

Halgren, T. A. 2009. Identifying and characterizing binding sites and assessing druggability. Journal of Chemical Information and Modeling 49: 377-389.

Halgren, T. A., R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard*, et al.* 2004. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. Journal of Medicinal Chemistry 47: 1750-1759.

Harder, E., W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang*, et al.* 2016. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. Journal of Chemical Theory and Computation 12: 281-296.

Harris, L. J., A. Saparno, A. Johnston, S. Prisic, M. Xu, S. Allard*, et al.* 2005. The maize *An2* gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. Plant Molecular Biology 59: 881-894.

Helliwell, C. A., J. A. Sullivan, R. M. Mould, J. C. Gray, W. J. Peacock, and E. S. Dennis. 2001. A plastid envelope location of *Arabidopsis ent*-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. The Plant Journal 28: 201-208.

Ishizawa, K. 2014. Intracellular pH regulation of plant cells under anaerobic conditions. Low-Oxygen Stress in Plants: 21:59-74. DOI: 10.1007/978-3-7091-1254-0_4

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8: 275-282.

Katoh, K., and D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30: 772-780.

Köksal, M., H. Hu, R. M. Coates, R. J. Peters, and D. W. Christianson. 2011. Structure and mechanism of the diterpene cyclase *ent*-copalyl diphosphate synthase. Nature Chemical Biology 7: 431-433.

Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547-1549.

Lawrence, C. J., Q. Dong, M. L. Polacco, T. E. Seigfried, and V. Brendel. 2004. MaizeGDB, the community database for maize genetics and genomics. Nucleic Acids Research 32: D393-D397.

Lemke, C., K. C. Potter, S. Schulte, and R. J. Peters. 2019. Conserved bases for the initial cyclase in gibberellin biosynthesis: from bacteria to plants. Biochemical Journal 476: 2607-2621.

Monaco, M. K., T. Z. Sen, P. D. Dharmawardhana, L. Ren, M. Schaeffer, S. Naithani*, et al.* 2013. Maize metabolic network construction and transcriptome analysis. The Plant Genome 6: 1-12. https://doi.org/10.3835/plantgenome2012.09.0025

Nagata, T., A. Hosaka-Sasaki, and S. Kikuchi. 2016. The evolutionary diversification of genes that encode transcription factor proteins in plants, pp. 73-97, Plant Transcription Factors. Elsevier.

Nelissen, H., B. Rymen, Y. Jikumaru, K. Demuynck, M. Van Lijsebettens, Y. Kamiya*, et al.* 2012. A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Current Biology 22: 1183-1187.

Olsson, M. H., C. R. Søndergaard, M. Rostkowski, and J. H. Jensen. 2011. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. Journal of Chemical Theory and Computation 7: 525-537.

Prisic, S., M. Xu, P. R. Wilderman, and R. J. Peters. 2004. Rice contains two disparate *ent*-copalyl diphosphate synthases with distinct metabolic functions. Plant Physiology 136: 4228-4236.

Sakamoto, T., K. Miura, H. Itoh, T. Tatsumi, M. Ueguchi-Tanaka, K. Ishiyama*, et al.* 2004. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiology 134: 1642-1653.

Schaefer, C., and B. Rost. 2012. Predict impact of single amino acid change upon protein structure. BMC Genomics 13: 1-10. https://doi.org/10.1186/1471-2164-13-S4-S4

Schmelz, E. A., F. Kaplan, A. Huffaker, N. J. Dafoe, M. M. Vaughan, X. Ni*, et al.* 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. Proceedings of the National Academy of Sciences 108: 5455-5460.

Su, P.-H., and Y.-H. Lai. 2017. A reliable and non-destructive method for monitoring the stromal pH in isolated chloroplasts using a fluorescent pH probe. Frontiers in Plant Science 8: 2079.

Sun, T. p., and Y. Kamiya. 1997. Regulation and cellular localization of *ent*-kaurene synthesis. Physiologia Plantarum 101: 701-708. https://doi.org/10.1111/j.1399-3054.1997.tb01054.x

Tastan Bishop, A. O., T. A. De Beer, and F. Joubert. 2008. Protein homology modelling and its use in South Africa: research in action. South African Journal of Science 104: 2-6.

Tholl, D. 2015. Biosynthesis and biological functions of terpenoids in plants. Biotechnology of Isoprenoids: 148:63-106. DOI: 10.1007/10_2014_295.

Toyomasu, T., M. Usui, C. Sugawara, Y. Kanno, A. Sakai, H. Takahashi*, et al.* 2015. Transcripts of two *ent*-copalyl diphosphate synthase genes differentially localize in rice plants according to their distinct biological roles. Journal of Experimental Botany 66: 369-376.

Vaughan, M. M., S. Christensen, E. A. Schmelz, A. Huffaker, H. J. Mcauslane, H. T. Alborn*, et al.* 2015. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. Plant, Cell & Environment 38: 2195-2207.

Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny*, et al.* 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research 46: W296-W303.

White, C. N., W. M. Proebsting, P. Hedden, and C. J. Rivin. 2000. Gibberellins and seed development in maize. I. Evidence that gibberellin/abscisic acid balance governs germination versus maturation pathways. Plant Physiology 122: 1081-1088.

Chapter 4: The role of kauralexin synthesis in maize defences against fall armyworm larvae

4.1 Introduction

Maize, *Zea mays*, is a major food source for millions around the globe, especially in African and Latin American countries (Shiferaw *et al.* 2011). One of the limiting factors for maize production in Africa, which accounts for up to 20% of the 54% annual yield loss, is attributed to animals and insects (Oerke 2006, Shiferaw *et al.* 2011). The increase of globalisation, trade and global climate has led to the further spread of insect pests (Bebber *et al.* 2014). Reducing maize yield losses to pathogens and insect pests through developing resistant maize varieties shows promise and could be key in increasing total yields and improving the stability of this sector with the challenge of insect pests into new regions (Shiferaw *et al.* 2011).

The fall armyworm, *Spodoptera fruigiperda,* is a lepidopteran insect pest of maize, which was introduced into Africa in 2016 and South Africa in 2017 (Goergen *et al.* 2016, Jacobs *et al.* 2018). The fall armyworm's introduction and spread in Africa is a concern with some estimates that the damage done to maize could amount to $3 billion USD in a period of 12 months (Stokstad 2017). The established control measures for this pest in Africa is limited, with appropriate chemical products and the knowledge to properly apply these products being scarce (Stokstad 2017). The use of *Bt* (*Bacillus thuringiensis*) maize and *Bt* based products have shown promising results in North and South America in controlling the fall armyworm and may be a good option for maize farmers in Africa, but it should be noted that resistance to some of these insecticidal proteins has already been noted (Huang *et al.* 2014, Botha *et al.* 2019).

Breeding resistant maize by identifying and incorporating native resistance from maize inbred lines and hybrids into conventional breeding programs is part of developing an integrated approach to managing this pest (Stokstad 2017). Kauralexins are terpenoid defence compounds naturally synthesized in maize (Schmelz *et al.* 2011, Tholl 2015, Block *et al.* 2019).

These compounds have been observed to express insect antifeedant activity to insect pests such as the European corn borer (Schmelz *et al.* 2011) as well as express antimicrobial properties against fungal pathogens, such as *Fusarium spp.* (Harris *et al.* 2005, Veenstra *et al.* 2019). The effect of kauralexins on fall armyworm larvae have however not been reported on in published literature.

The aim of this study is to investigate the effect of kauralexins on fall armyworm larval feeding and feeding preference. The effect of these compounds on fall armyworm larvae can then be used to determine whether this defence pathway plays a critical role in the innate maize defences against this insect pest and if this pathway should be considered when breeding for resistant maize lines. To determine the effect of these compounds on fall armyworm larvae feeding and feeding preference, non-choice and choice herbivory assays were performed using maize which synthesizes these compounds (W22 maize inbred line) and maize which synthesizes low to no kauralexins due to a gene mutation in a kauralexin biosynthesis pathway enzyme (An2 mutant maize line).

4.2 Methods

4.2.1 Biological materials

Two maize inbred line were used in this study, *Zea mays* W22 and an *An2* mutant maize line. Fall armyworm, *Spodoptera frugipera*, larvae were used in the various feeding assays. The W22 maize line contained a functional An2 protein (Zm00004b001917; maizeGDB protein accession number) while the An2 mutant included a transposon in the *An2* gene resulting in a gene knockout for the An2 protein (Vaughan *et al.* 2015). Seed for both maize lines, W22 and mutant An2 maize, were acquired from the University of California San Diego. The seed was grown to maturity and plants self-pollinated at the University of Pretoria experimental farm to produce more seed. The fall armyworm larvae used, were originally collected from the University of Pretoria's experimental farm as well as from the Agricultural Research Council (ARC) in Roodeplaat. The strain identity of the fall armyworm colony was determined by amplifying and sequencing the mitochondrial *cytochrome oxidase I* (*COI*) (Goergen *et al.* 2016, Nagoshi *et al.* 2017) and the nuclear *triosephoshate isomerase* (*Tpi*) (Nagoshi *et al.* 2017) genes (results are presented in chapter 2 of this study).

4.2.2 Growing maize plantlets

W22 and *An2* mutant maize plants were grown in climate controlled rooms in the Department of Zoology and Entomology at the University of Pretoria (Figure 4.1). Plants were grown in individual pots at constant 25°C, with 20 - 22% relative humidity and fluorescent light intensity of 24.29 – 35.08 µmol / s / m^2 (1800 – 2600 lux, 12h:12h, L:D cycle) for approximately 3 weeks (growth stage V2). The growth medium was an outdoor potting soil (Bark Unlimited©) and plants were watered tri-weekly, or as needed to ensure well-watered conditions, using a

fertilizer solution comprised of water with 1.28 g/2L hygroponic and 0.98g/2L solu-cal (Hygrotech, Pretoria, South Africa).

Figure 4.1: Maize inbred line W22 and mutant An2 plants grown in a climate controlled room in the Department of Zoology and Entomology, University of Pretoria. Growth conditions were constant 25°C, with 20 - 22% relative humidity and a light intensity of 24.29 – 35.08 µmol / s / m^2 (1800 – 2600 lux) using fluorescent light (12h:12h, L:D cycle).

4.2.3 DNA extraction of maize samples

DNA extraction was performed from maize leaves to test for the presence / absence of the An2 mutation in maize plants used for the feeding experiments. Two maize leaves were harvested from each of the plants, flash-frozen in liquid N_2 and stored at -80°C. The leaf samples were ground up using a mortar and pestle. A Cetyl-Trimethyl-Ammonium-Bromide (CTAB) based DNA extraction method was used to extract the maize gDNA. The particular version of the CTAB protocol was adapted from a research article, modifications by Dr DL. Nsibo from the University of Pretoria were included (Wu *et al.* 2001). The CTAB buffer (2% CTAB, 200mM Tris-HCl, 200mM EDTA, 2.5mM NaCl) was made to a pH of 8.0 and autoclaved. Activated CTAB buffer was made by pipetting 0.5 mL of the CTAB buffer into a sterile container, adding 0.02g of polyvinylpyrrolidone (PVP) and 2.5 µL of β-mercaptoethanol,

this was scaled up to the final amount needed while maintaining the respective ratios. Approximately 200 mg of ground leaf material was transferred to liquid N_2 cooled Eppendorf tubes, 800 µL of the activated CTAB buffer was added to each sample. Samples were vortexed for 5 min and incubated at 65°C for one hour. The samples were mixed by inverting every 30 min. After incubation 800 µL of chloroform was added to each sample, the samples were mixed and centrifuged for 12 min at 11 337 x g. Approximately 600 µL of the supernatant was transferred into new 2 ml Eppendorf tubes, following which 500 units of DNA free RNase A (Norgen Biotek Corp., Ontario, Canada) were added to each sample and incubated for 30 min at 37°C. Once incubated, 600 µL of a phenol-chloroform solution (1:1) was added to each sample and vortexed for 1 minute. Samples were centrifuged for 12 min at 11 337 x g and 500 µL of the top aqueous solution was transferred to new 2 ml Eppendorf tubes. Pre-chilled isopropanol was added to each sample and incubated overnight at -20°C to precipitate the DNA. Samples were centrifuged for 15 min at 11 337 x g and the supernatant decanted. The pelleted DNA was washed by adding 1 ml pre-chilled 70% ethanol and centrifuging for 15 min at 11 337 x g, decanting the supernatant and this wash step was repeated. The pelleted DNA was airdried overnight and resuspended in 50 µL of ddH_2O. The quality and concentration of the extracted DNA was evaluated on a NanoDrop™ 2000 Spectrophotometer (Thermo Scientific™, Massachusetts, United States) and by visualising on a 1% TAE agarose gel (80 V, 400 mA, 50 min) to ensure no major DNA degradation had taken place that could adversely impact PCR screening. The DNA was stored at -20°C.

4.2.4 Maize An2 gene screening and sequencing

Maize DNA was screened to determine the presence or absence of the mutant *An2* gene. To identify the presence of the mutation, a primer pair which spans a portion of the *An2* gene and the transposable element (*Ds* insertion) which results in the gene knockout is used. This primer pair for mutation identification is F312 (5'-GATCGCCTGGAGCGTCTCGG-3') and JGP3 (5'-ACCCGACCGGATCGTATCGG-3') (Inqaba Biotech, Pretoria, South Africa), which

forms a 410 bp PCR product when the mutation is present and no PCR product when it is absent. To identify the presence of the functional *An2* gene a primer pair in which both primers bind to regions of the *An2* gene are used. These primers identify functional *An2* by spanning the gene region in which the mutation inducing transposable element would be present if it were inserted (not present in W22). The primer pair for functional *An2* identification is F312 and *ZmAn2* 1815R (5'-CGTCACCGCCTGGTTCGACT-3') (Whitehead Scientific, Cape Town, South Africa), which forms a 535 bp PCR product when the functional *An2* gene is present and no PCR product when the mutation is present. The reverse primer ZmAn2 1815R is a new primer developed by Prof Dave Berger at the University of Pretoria. Presence of both PCR products indicated heterozygosity, the presence of both a mutant and functional *An2* gene copies. A positive PCR control was included, which involved amplifying an actin gene using the Banana actin primers, B. actin F (5'-ACCGAAGCCCCTCTTAACCC-3') and B. actin R (5'-GTATGGCTGACACCATCACC-3') (Inqaba Biotech, Pretoria South Africa) (Van den Berg *et al.* 2004). The PCR product was the result of amplification with the Banana actin F and R primers (B. actin F and B. actin R) corresponding to an ~200 bp actin gene region. In total, 23 *An2* mutant and 11 W22 maize plants were screened by PCR amplification of the mutant *An2* and functional *An2* gene across both the non-choice and choice herbivory assays. Polymerase chain reaction (PCR) was performed in a 12 µL reaction mixture containing 6.8 µL ddH_2O, 0.6 µL forward primer (10 µM), 0.6 µL reverse primer (10 µM), 3 µL Taq 2X Master Mix Red (Amplicon Inc., Odense, Denmark) and 1 µL genomic DNA (approximately 50 ng). Samples were placed in a 2720 Thermal Cycler (Applied Biosystems, Foster City, California, United States) and programmed with either of the two cycling conditions described below depending on the primer pair used. The F312 and JGP3 (mutation identification) cycling conditions are as follows: 94°C (5 min), followed by 30 cycles of 94°C (30 s), 59°C (30 s), 72°C (30 s), and a final segment of 72°C (30 s). The F312 and *ZmAn2* 1815R (functional *An2* identification) cycling conditions are as follows: 94°C (5 min), followed by 30 cycles of 94°C (30 s), 62°C (1 min), 72°C (2 min), and a final segment of 72°C (2 min). The B. actin F and B. actin R cycling conditions are as follows: 94°C (2 min), followed by 30 cycles of 94°C (15 s),

58.3°C (15 s), 72°C (20 s), and a final segment of 72°C (20 s). The PCR products were visualised by gel electrophoresis on a 1% TAE agarose gel (80 V, 400 mA, 50 min) with the GeneRuler 100 bp DNA ladder (Thermo Fischer Scientific™, Massachusetts, United States). PCR products were cleaned using ExoSAP-IT (Thermo Fischer Scientific™, Massachusetts, United States). Exosap was added to each of the PCR product samples, 2 µL Exosap for every 5 µL of PCR product, and the following incubation steps were utilised: 37°C (15 min) and 80°C (15 min). The sanger sequencing reactions were performed in a 10 µL reaction mixture containing 2 µL BigDye® (Thermo Fischer Scientific™, Massachusetts, United States), 1 µL 5X sequencing buffer, 0.5 µL of either forward or reverse primer (10 µM), 2.5 µL purified PCR product and 4 µL ddH_2O. The cycling conditions used were as follows: 94°C (1 min), followed by 25 cycles of 96°C (10 s), 56°C (5 s) and 60°C (4 min). The samples were then sequenced by the sequencing facility of the University of Pretoria. The PCR products of interest for both screening PCR's were sequenced and confirmed to be the correct gene (results not shown).

4.2.7 Non-choice herbivory assays

To determine whether higher levels of kauralexins negatively affected fall armyworm feeding and weight gain, non-choice feeding assays were performed. Fall armyworm larvae of the third instar were used (Figure 4.2 A). The non-choice herbivory assays consisted of experimental replicates and negative controls. The experimental replicates involved introducing individual fall armyworm larvae onto either W22 or An2 mutant maize plants, keeping the individual larvae on specific maize leaves using 90 mm diameter leaf traps (Figure 4.2 B). The negative controls were maize plants which had a leaf trap attached without any larvae inside, this was to determine whether the leaf traps caused any leaf damage. The first no-choice herbivory assay experimental replicates consisted of 13 An2 mutant replicates and 14 W22 replicates. The second non-choice herbivory assay experimental replicates consisted of 14 An2 mutant replicates and 14 W22 replicates. The larvae were only allowed to feed on the assigned leaf

within the leaf trap. Larvae were weighed prior to introduction and after the feeding period to determine weight gain using an AR0640 Adventurer OHAUS scale (Ohaus corp, Pine Brook, NJ USA). Photographs were taken of the maize leaves before and after the feeding period using a Nikon 90D DSLR camera. The camera was mounted on a tripod and the leaves on a platform to ensure consistency between photographs. These photographs were then used to determine the leaf area consumed (process explained in section 4.2.9). The larvae were allowed to feed for 18h and all replicates were kept at constant conditions (constant 25°C; approximately 10.25 µmol / s / m^2; 12h:12h, L:D cycle and 20% relative humidity) in IncoCool (Labotec) and Labcon growth chambers (Labex) (Figure 4.2 B).

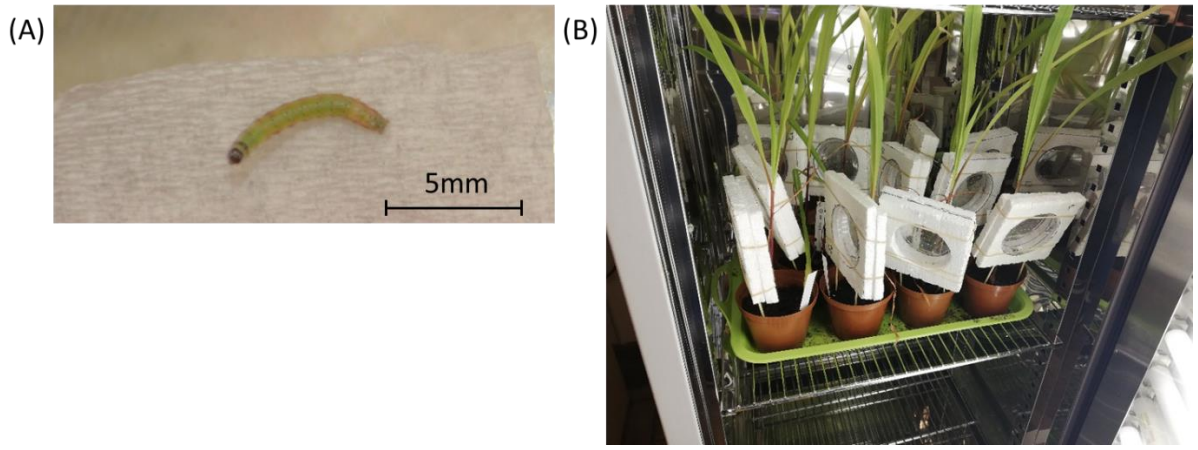


Figure 4.2: Non-choice herbivory assay. (A) third instar fall armyworm larvae used in non-choice herbivory assays. Scale bar represents 5 mm for size determination. (B) Non-choice herbivory assay replicates with leaf traps containing fall armyworm larvae on the leaves in a climate controlled growth chamber (constant 25°C; approximately 10.25 µmol / s / m^2; 12h:12h, L:D cycle and 20% relative humidity).

4.2.8 Choice herbivory assays

To determine whether kauralexins influence fall armyworm larvae feeding preference, choice herbivory assays were performed using W22 and mutant An2 maize. Fall armyworm larvae of the second instar were used for these experiments (Figure 4.3 A). The choice herbivory assays consisted of mixed genotype pair replicates (also referred to as mixed pairs), same genotype pairs (also referred to as same pairs) and negative control pairs. The mixed pair replicates consist of one of each of the maize genotypes, one W22 maize plant and one mutant

An2 maize plant. The same pair replicates consist of a pair of maize plants, the individuals either being both W22 or mutant An2 maize plants. The negative controls consisted of a mixture of genotype pairs with leaf traps containing no larvae. The first choice herbivory assay included 16 mixed pair replicates, 9 An2 mutant same pairs and 6 W22 same pairs. The second choice herbivory assay included a total of 21 mixed pair replicates, 10 An2 mutant same pair replicates and 11 W22 same pair replicates.

Single larvae were contained in leaf traps (20 mm diameter) clipped onto two leaves each (Figure 4.3 B), one leaf from either maize pair individual, allowing the larvae the ability to choose between feeding on either individual of the pair. The same pair replicates also served as a type of non-choice assay, as there was no choice available of a different genotype. The negative controls were included to account for any unexpected leaf damage or area loss. The larvae were weighed prior to introduction and after the feeding period to determine weight gain an AR0640 Adventurer OHAUS scale. Maize leaves were photographed in their respective pairs after the feeding period (Figure 4.3 C). The larvae were allowed to feed for 48h and the experiment was conducted in a climate controlled room (constant 25°C; approximately 24.29 – 35.08 µmol / s / m^2; 12h:12h, L:D cycle and 20% relative humidity).

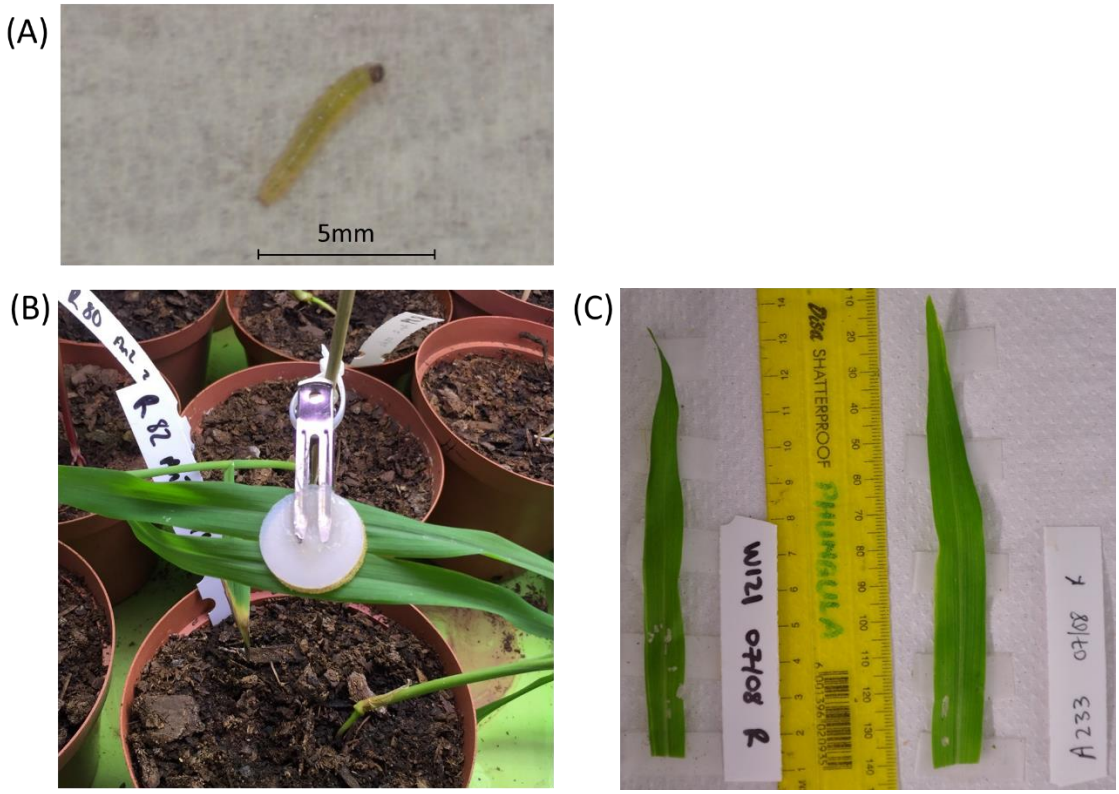


Figure 4.3: Choice herbivory assay replicates. (A) Second instar fall armyworm larvae use in choice herbivory assays. (B) Small leaf trap containing a fall armyworm larvae onto a choice herbivory assay replicate pair. (C) Photograph of W22 (left) and mutant An2 (right) maize leaves from a replicate pair post feeding period.

4.2.9 Digital Image Analysis

Images of the maize leaves from non-choice and choice herbivory assays were analysed using ImageJ (Schneider *et al.* 2012). The scale was determined and set using a ruler included in the images. Images from non-choice herbivory assays of the maize leaves before and after feeding were used to determine leaf area consumed. Images were converted to 8-bit images, the threshold altered to highlight the entire leaf area and the measure function used to determine the leaf area. The final leaf area after feeding was subtracted from the leaf area prior larval introduction to determine the leaf area consumed. Images from choice herbivory assays of the maize leaves after feeding were used to determine leaf area consumed. Feeding areas were outlined using the freehand function which highlighted the missing area and the

measure function was used to determine the leaf area consumed. The total leaf area was tallied for each of the maize plants.

4.2.10 Statistical analysis

The non-choice and choice herbivory assays were comprised of larval mass gain and leaf area consumed data. Larval mass gain data was transformed into relative mass gain data by dividing the mass gain by the initial mass of the larvae, this was to correct for the differences in initial mass. A Shapiro-Wilk test was conducted on the separate data sets to determine whether the data was normally distributed. Non-normal data was transformed by $\log_{10}$. Once the data was confirmed to be normally distributed, an independent samples t-test was conducted on non-choice larval relative mass gain data, non-choice leaf area consumed data and the choice leaf area consumed data of the same pair replicates. The independent t-tests were to determine if the difference in means between the two genotypes were significant. One-way analysis of variance (ANOVA) was performed on choice larval relative mass gain data at a 95% confidence level to determine whether there was any significant differences in relative mass gain of larvae feeding on the different pairs of maize genotypes. The choice leaf area consumed data of mixed pair replicates is paired data and thus a paired t-test was performed at a 95% confidence level to determine whether there was any significant difference in feeding on either of the two maize genotypes. These analyses were performed in the open source software R version 3.6.1 (www.r-project.org) in conjunction with the graphical user interface package R Commander (Fox 2005).

4.3 Results

4.3.1 Molecular screening of An2 in maize plants

Maize leaves collected from plants used in the non-choice herbivory assays and the choice herbivory assays were screened to detect the presence of the mutant and functional *An2* gene in the W22 and An2 mutant plants. Representative samples were collected from each herbivory assay and screened. The *An2* PCR screening for maize samples from choice herbivory assay 2 is presented below in Figure 4.4, while the gel electrophoresis images for the remaining herbivory assays can be found in the supplementary information (Supplementary figure 4.1 and 4.2).

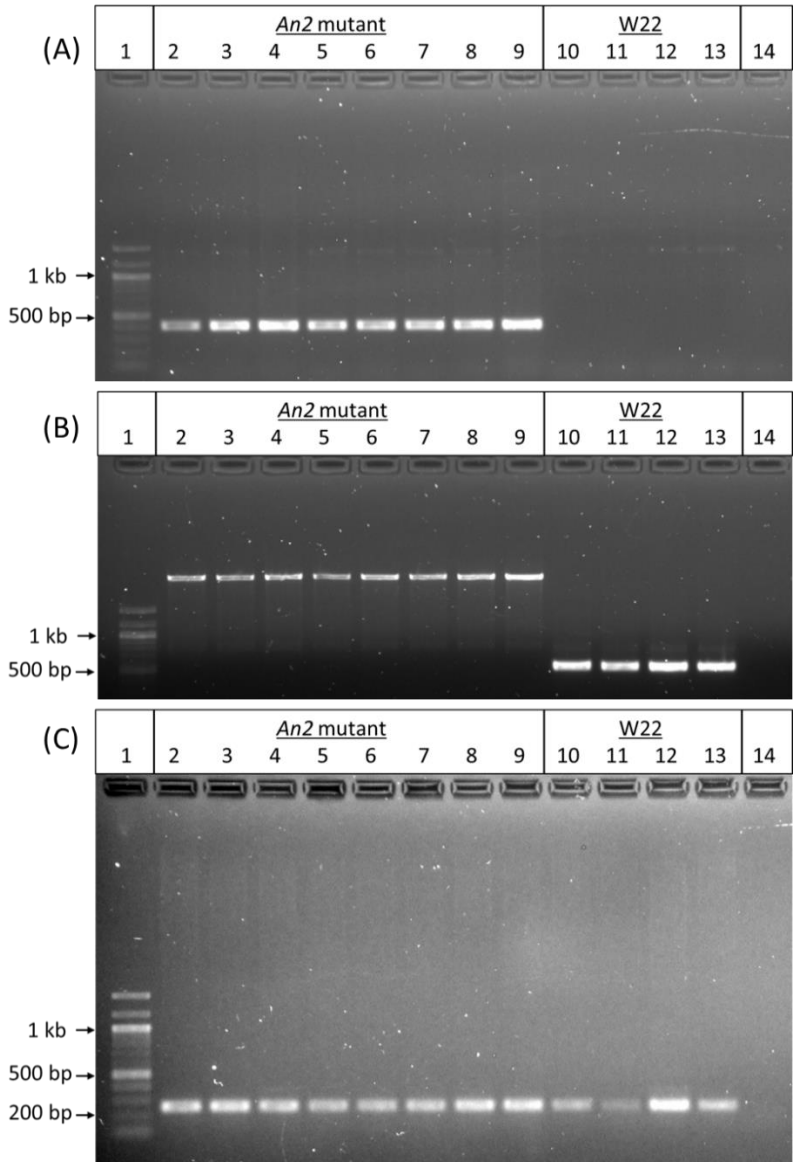


Figure 4.4: Maize *An2* screening of maize samples collected from choice herbivory assay 2. (A) Mutant *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and JGP3 primers). (B) Functional *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and 1815R primers). (C) Positive control PCR using banana actin primers of mutant *An2* and W22 maize (B.actin F and B.actin R primers). Lanes 2 – 9: An2 mutant maize; Lanes 10 – 13: W22 maize; Lane 14: non-template/water control; Lane 1: 100 bp DNA Ladder (New England Biolabs). PCR products were resolved on a 1% TAE agarose gel subjected to 80V for 50 min with ethidium bromide.

In Figure 4.4 (A) a 410 bp PCR product is observed in lanes 2 – 9 of the mutant *An2* maize samples. The 410 bp PCR product indicates presence of the mutant *An2* gene (*Ds* insertion present resulting in gene knockout). This PCR product is not observed in lanes 10 – 13 of the W22 maize samples, which indicates that no mutant *An2* gene was present. Lane 14 was the

non-template control, no signs of amplification can be observed. In Figure 4.4 (B) a 535 bp PCR product is observed in lanes 10 - 13 of the W22 maize samples. The 535 bp PCR product indicates the presence of the functional *An2* gene (no *Ds* insertion present). This PCR product is not observed in lanes 2 – 9 of the mutant *An2* maize samples, but a faint PCR product > 1.5 Kb is observed. Lane 14 was the non-template control, no signs of amplification can be observed. Observing the results from Figure 4.4 (A) and (B) together indcates that the mutant *An2* and W22 maize samples were homozygous for either the *An2* mutant gene (mutant *An2* maize, lanes 2 – 9) or the functional *An2* gene (W22 maize, lanes 10 – 13). In Figure 4.4 (C) a 200 bp PCR product is observed in lanes 2 – 13 of both mutant *An2* and W22 maize samples, this is the product of the *B. actin* primers used as controls to indicate good quality gDNA. A summary of the *An2* screening PCR's for all mutant *An2* and W22 maize samples collected from the non-choice and choice experiments illustrated below in Table 4.1.

Table 4.1: Summary of *An2* screening results of maize samples collected from the non-choice and choice feeding assays.

mutant *An2* maize

Experiment	n	Hom *An2* mutant*	Het *An2* mutant*	Hom functional *An2**
Non-choice experiments 1 and 2	8	2	4	2
Choice experiment 1	7	2	3	2
Choice experiment 2	8	8	0	0

W22 maize

Experiment	n	Hom *An2* mutant*	Het *An2* mutant*	Hom functional *An2**
Non-choice experiments 1 and 2	4	0	0	4
Choice experiment 1	3	0	0	3
Choice experiment 2	4	0	0	4

**Hom indicates homozygous and Het indicates heterozygous.*

The *An2* mutant maize samples collected from the non-choice experiments and choice experiment 1 were homozygous and heterozygous for the *An2* mutant gene (11 out of 15 samples), while 4 samples did not have the mutant *An2* gene. The *An2* mutant maize samples collected from choice herbivory assay 2 were all homozygous for the *An2* mutant gene, as

noted also in Figure 4.2. The W22 maize samples collected from the non-choice and choice experiments were all homozygous for the function *An2* gene, no mutant *An2* genes were present out of the total 11 samples screened. The *An2* screening gel electrophoresis images of the non-choice and choice experiment 1 maize samples are included in (Supplementary Figure 4.1 and 4.2).

4.3.2 Non-choice herbivory assay

The non-choice herbivory assays were performed using third instar fall armyworm larvae on W22 and An2 mutant maize plants. The average total leaf area consumed in non-choice experiment 1 on both genotypes was between 3 and 3,5 cm^2 and in non-choice experiment 2 on both genotypes was between 1,75 and 2 cm^2 (Figure 4.5).The leaf area consumed data in both non-choice herbivory assays were not significantly different between the maize genotypes. The negative controls did not exhibit any leaf damage.

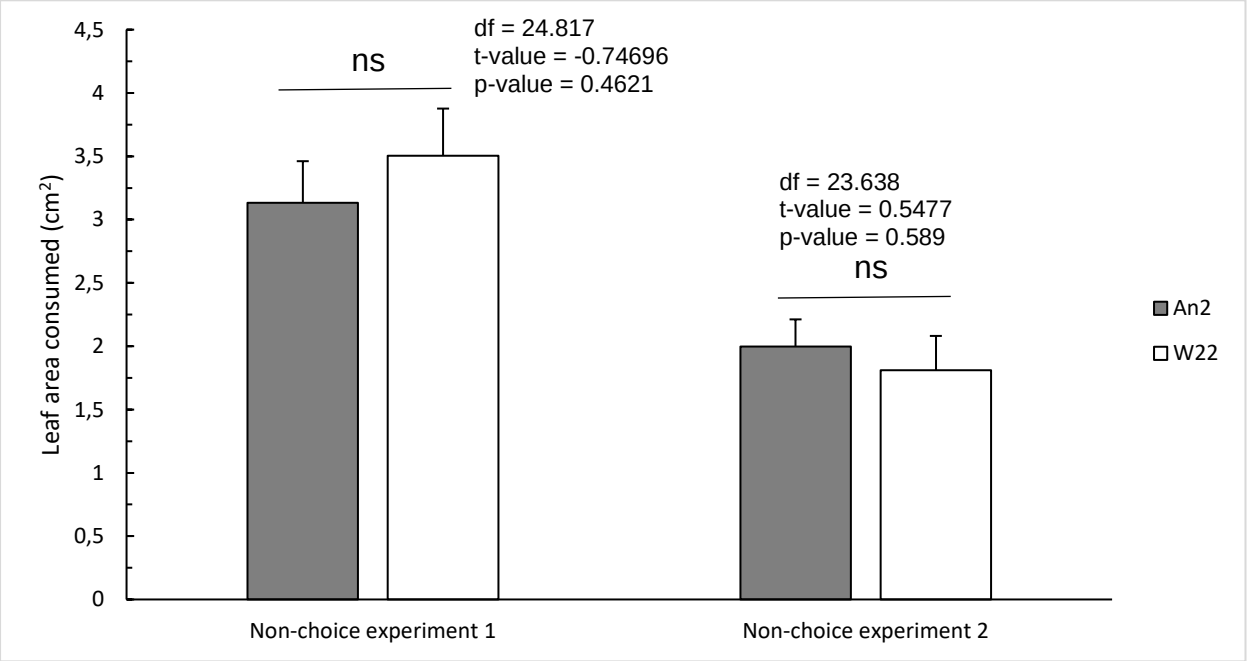


Figure 4.5: The mean leaf area consumed (cm^2) (+ SEM) by instar fall armyworm larvae on the maize genotypes An2 and W22 after an 18hr feeding period. For experiment 1, n = 13 An2 and n = 14 W22. For experiment 2, n = 15 An2 and n = 13 W22. Feeding occurred under 20% humidity, 12h / 12h light-dark cycles and constant 25°C. Non significance is denoted by ns and a line above the compared genotypes, this was determined by performing an independent t-test.

The relative larval mass gain of the larvae in the first non-choice herbivory assay gained an average relative mass of approximately 1,3 mg for larvae feeding on either of the genotypes. The relative larval mass gain of the larvae in the second non-choice herbivory assay gained an average relative mass of approximately 0,9 mg for larvae feeding on either of the genotypes. There was no significant difference in relative mass gain between larvae which fed on either of the maize genotypes (Figure 4.6).

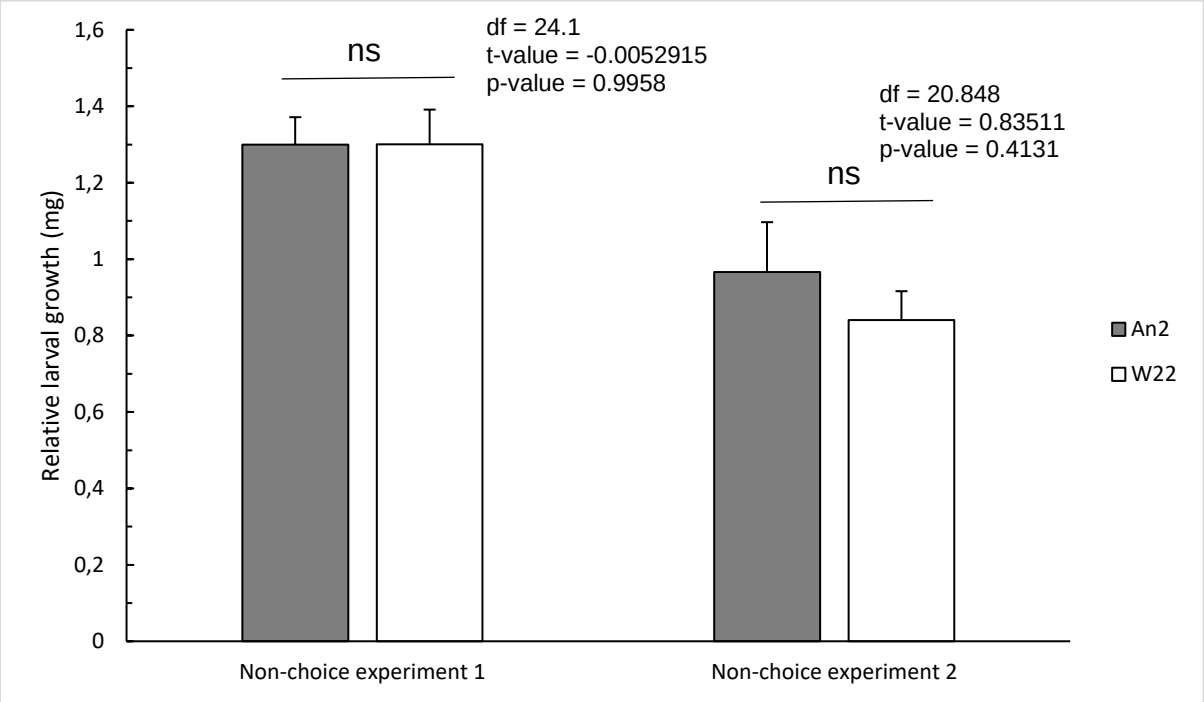


Figure 4.6: The mean relative larval growth (mg) (+ SEM) by third instar fall armyworm larvae on the maize genotypes An2 and W22 after an 18hr feeding period. Single fall armyworm caterpillars were isolated onto their own respective leaves of either maize genotype. For experiment 1, n = 13 An2 and n = 14 W22. For experiment 2, n = 14 An2 and n = 14 W22. Feeding occurred under 20% humidity, 12h / 12h light-dark cycles and constant 25°C. Non significance is denoted by ns and a line above the compared genotypes, this was determined by performing an independent t-test.

4.3.3 Choice herbivory assay

The choice herbivory assays were performed to determine if kauralexins influence larval feeding preference. The leaf area consumed in the mixed pairs (one of each genotype) was determined to not be significantly different between the two genotypes (Figure 4.7). In choice

herbivory assay 1 the mixed pair replicates exhibited an average of approximately 0,5 cm^2 leaf area consumed for either genotype and in choice herbivory assay 2 the mixed pair replicates exhibited an average of approximately 0,6 cm^2 leaf area consumed for either genotype. The same pair replicates also did not exhibit any significant difference in leaf area consumed between the two genotypes. In the same pair replicates of choice herbivory assay 1, the average leaf area consumed was approximately between 1,3 and 1,4 cm^2 for either genotype and in choice herbivory assay 2, the average leaf area consumed was approximately between 0,9 and 1,3 cm^2 for either genotype. These same pair replicates served as a second set of "non-choice" herbivory assays.

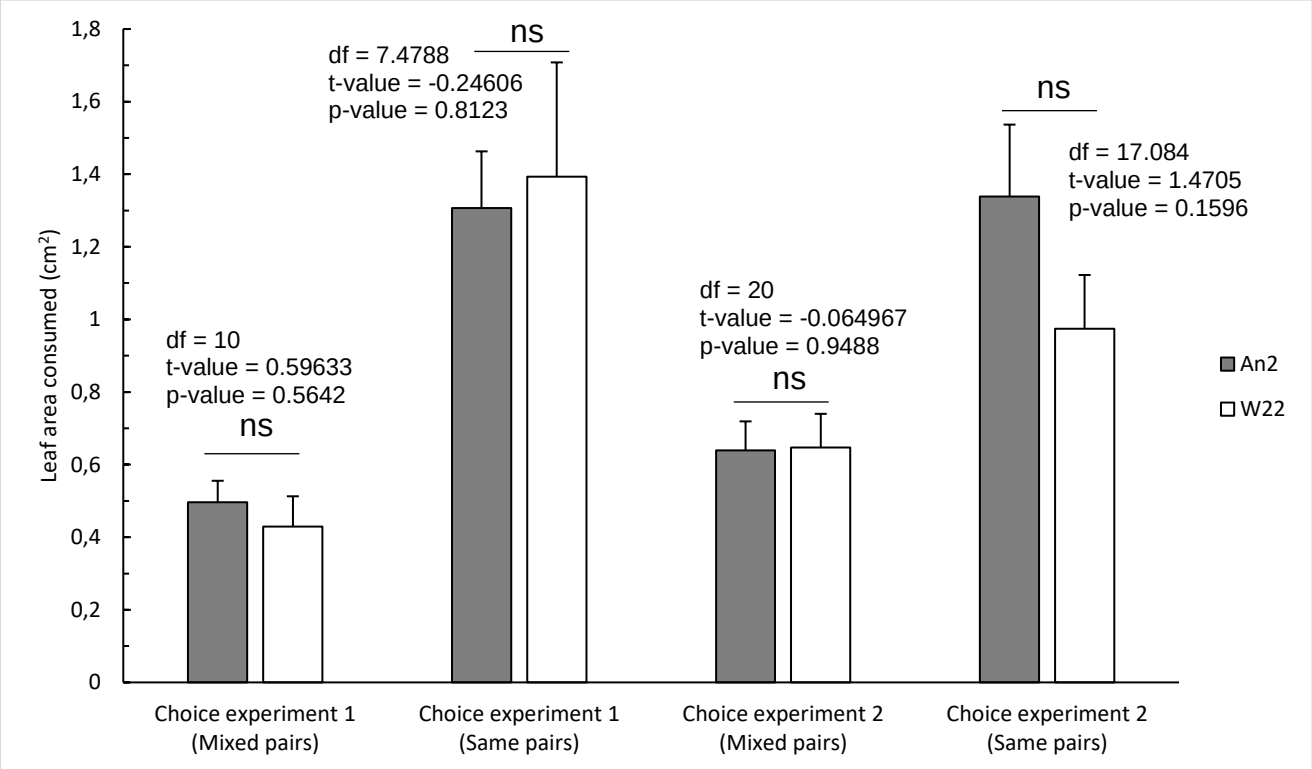


Figure 4.7: The mean leaf area consumed (cm^2) (+ SEM) by second instar fall armyworm larvae on An2 mutant and W22 maize in the choice herbivory experiments. Choice experiments were comprised of mixed genotype and same genotype pairs with larvae feeding over a 48hr period. (Choice experiment 1 (Mixed pairs): n = 16; Choice experiment 1 (Same pairs): An2 n = 9, W22 n = 6; choice experiment 2 (Mixed pairs): n = 21; Choice experiment 2 (Same pairs): An2 n = 10, W22 n = 11) Feeding occurred under 20% humidity, 12h / 12h light-dark cycles and constant 25°C. Non significance is denoted by ns and a line above the compared genotypes, this was determined using a dependent t-test (mixed pairs) and an independent t-test (same pairs).

The relative larval mass gain of larvae in the first choice herbivory assay on the mixed and same pair replicates were approximately 5 mg. The relative larval mass gain of larvae in the second choice herbivory assay on the mixed and same pair replicates were approximately 4 mg. There was determined to be no significant difference in relative mass gain of larvae feeding on either of the genotypes in both choice herbivory assay same pair replicates (Figure 4.8).

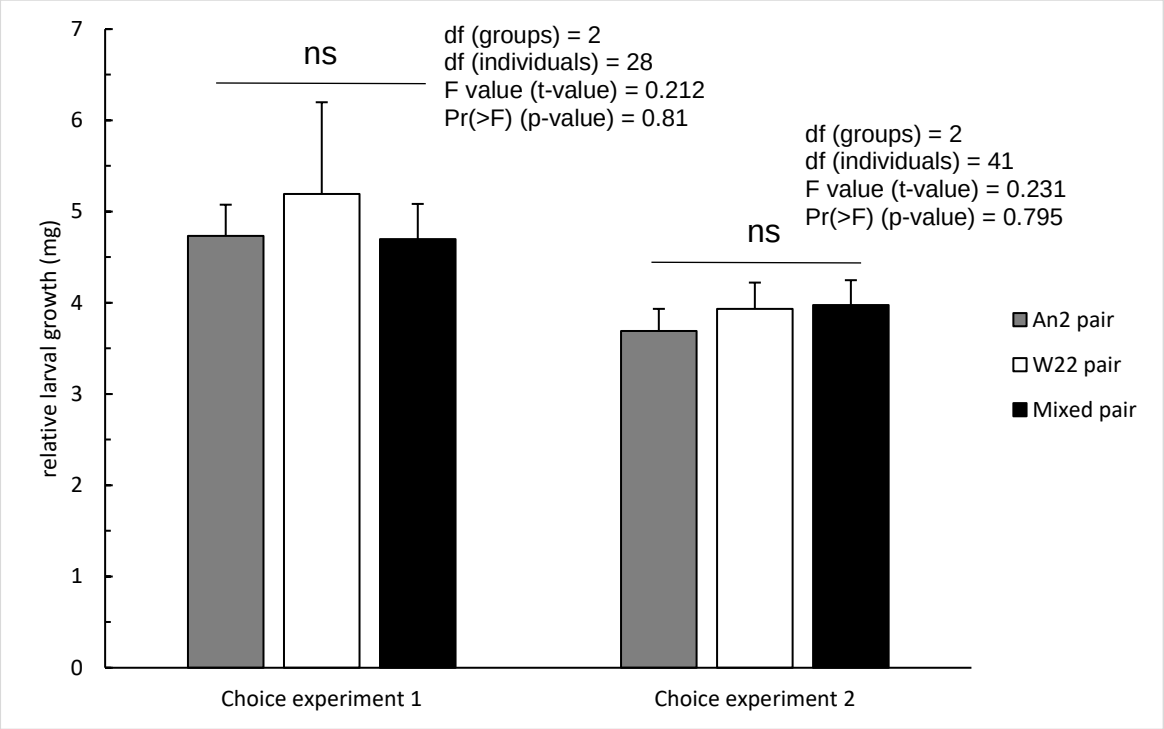


Figure 4.8: The mean relative larval growth (mg) (+ SEM) of second instar fall armyworm larvae on maize pairs of choice experiments 1 and 2 after an 48hr feeding period. (Choice experiment 1: An2 pairs n = 9, W22 pairs n = 6, Mixed pairs n = 16; Choice experiment 2: An2 pairs n = 10, W22 pairs n = 11, Mixed pairs n = 23). Feeding occurred under 20% humidity, 12h / 12h light-dark cycles and constant 25°C. Non significance is denoted by ns and a line above the compared genotypes, this was determined using an one-way ANOVA.

4.4 Discussion

Kauralexins are terpenoid secondary metabolites synthesized in maize that have been shown to play a significant role in maize defences against fungal pathogens and insect pests (Harris *et al.* 2005, Schmelz *et al.* 2011, Meyer *et al.* 2017, Veenstra *et al.* 2019). The fall armyworm, *S. frugiperda*, is a notorious maize insect pest which was recently introduced in South Africa in 2017 (Jacobs *et al.* 2018, Nagoshi *et al.* 2019). This pest is being controlled by using GMO *Bt* maize and various harmful chemical pesticides (Day *et al.* 2017), but many reports have shown an increase in resistance of the pest to these control measures (Huang *et al.* 2014). Investigating how innate maize defences interact with insect pests and how insect pests respond to specific plant defences may reveal which defence pathways are important in maize resistance to specific insect pest species.

In this study the interaction of fall armyworm larvae with kauralexin defence compounds in maize were investigated by performing non-choice and choice herbivory assays. The non-choice herbivory assays were performed to investigate the effect of kauralexins on fall armyworm larvae feeding. The differences in feeding between the maize genotypes were not significant, but there was a significant difference in leaf area consumed between the two experiments. The two experiments differed significantly in leaf area consumed because the fall armyworm larvae in the second experiment, although of the same instar (third instar), were younger by a few days and weighed less compared to the larvae in the first experiment. The amount of plant material consumed by larvae increases exponentially as larvae progress through the instar stages, with the final instar stage feeding more than all the previous instar stages combined (Luginbill 1928). The differences in relative larval mass gain of larvae feed on either of the maize genotypes was non-significant, but differences in relative larval mass gain between the experiments were significant for the same reason mentioned above. Although the datasets for the two experiments could not be merged, the final result for both

was the same in that kauralexins did not influence third instar fall armyworm feeding in the non-choice herbivory assays.

The larval instar used in the non-choice experiments was investigated as a potential reason for no differences being observed in larval feeding and mass gain on the two maize genotypes. The use of third instar larvae was predominantly due to practicality, these larvae were larger, less likely to die from handling in the experiments and fed more compared to neonates which are fragile and produce feeding windows in the maize leaves. Earlier instar larvae may be more susceptible to plant defence compounds compared to later instars (Zalucki *et al.* 2002). In the lepidopteran Monarch butterfly, *Danaus plexippus,* the fourth and fifth instar larvae were determined to be unaffected by the cardiac glycosides present in milkweed (Erickson 1973), but a rise in mortality was noted when first instar larvae were used (Zalucki and Brower 1992, Zalucki *et al.* 2001). Studies on the efficiency of certain pesticides and *Bt* toxins against *Spodoptera* species found similar results. When a *Bt* based spray was applied to second and fourth instar *Spodoptera littoralis* larvae, the earlier instar exhibited 20% mortality when the manufacturer's recommended LC_{50} was applied, while the same concentration did not result in any mortality of the later instar larvae (Valadez-Lira *et al.* 2012). Sixth instar fall armyworm larvae were found to be over 100 fold more tolerant to the pesticides methomyl, diazinon and permethrin when compared to third instar larvae (Yu 1983). Previous studies in which kauralexins were investigated for their defensive properties against the European corn borer, fifth / final instar larvae were used (Dafoe *et al.* 2011, Schmelz *et al.* 2011). One could argue that if these experiments were repeated using earlier instars or neonates, the results may be different and kauralexins may play a significant role in the maize defences against the larvae. An alternative may be to perform these experiments over longer periods of time and by allowing moths to lay eggs on the plants of interest in order to investigate the role of the plant defences on the most susceptible stage of the insect's life cycle, the neonate stage.

The choice herbivory assays were performed to investigate the effect of kauralexins on second instar fall armyworm feeding preference. The reason for the large difference in leaf area consumed in the mixed and same pair replicates is because the total leaf area consumed in the mixed pair replicates is split between two genotypes, while in the same pair replicates is only on one genotype. This means that to compare total leaf area consumed in choice replicates one would need to add the W22 and mutant *An2* maize leaf area consumed to be directly comparable to the same pair replicates. The same pair replicates were a type of non-choice replicates (pair comprised of two of the same genotype) included in the choice experiments to replicate the non-choice experiments using smaller larvae. There was no significant preference of one genotype over another in the choice experiments and no significant differences in feeding between genotypes in the same pair replicates. There were significant differences in leaf area consumed between experiments due to the second choice experiment using younger second instar larvae. There was also no significant difference in larval mass gain of larvae feeding on either genotype. Kauralexins did not affect the feeding preference of second instar fall armyworm larvae in the choice experiments. Use of the earlier instar (second instar) in the choice herbivory assays, which was significantly smaller in mass compared to the third instar larvae, did not make a difference to the final result.

The majority of the representative An2 mutant maize samples screened varied in homozygosity and heterozygosity for the mutation. This result was investigated as the potential reason for no differences in larval feeding being observed on the two different maize genotypes in either non-choice and choice herbivory assays. Kauralexin synthesis can be linked to the *An2* gene as previous studies have shown a knockout of this gene results in almost complete loss of kauralexin synthesis (Vaughan *et al.* 2015). Heterozygosity for this mutation still results in significantly less kauralexin synthesis compared to maize not containing the mutation (Wighard 2017). This means that there would be clear differences in terms of larval feeding compared to W22 maize if kauralexins played a critical role against the fall armyworm. Some exceptions were present however in the An2 mutant maize samples

screened which did not have the mutation at all. The An2 mutant maize plants without the mutation were in the minority (~26% of representative sample set) and unlikely to dilute the final results in such a way that no differences are noticed between genotypes. The reason for this variation presence and absences of the mutation may be due to the parental lines being heterozygous for the mutation. Mutant An2 maize heterozygous for the An2 mutation was most likely produced by accidental crossing of W22 and mutant An2 maize in the process of growing maize plants for the purpose of generating additional seed. In choice herbivory assay 2, all mutant An2 maize tested homozygous for the mutant *An2* gene, this seed was generated from new crosses using parental maize known to be homozygous for the mutant *An2* gene.

The screening performed to determine the presence and absence of the mutant *An2* gene was confirmed to amplify the correct gene regions through sequencing the PCR products. There was however some additional, unexpected PCR products observed when performing gel electrophoresis. The faint PCR product observed between 1200 and 1500 across all samples is most likely due to nonspecific binding. The PCR product >1.5 Kb observed in the mutant An2 samples is most likely the result of the forward and reverse primers (F312 and 1815R) binding before and after the *Ds* insertion site on the *An2* gene. In the case of the mutant *An2* gene in which the 1.8 Kb *Ds* insertion is present, the PCR would result in a >2 Kb PCR product. These unexpected PCR products do not take away from the reliability of the screening results.

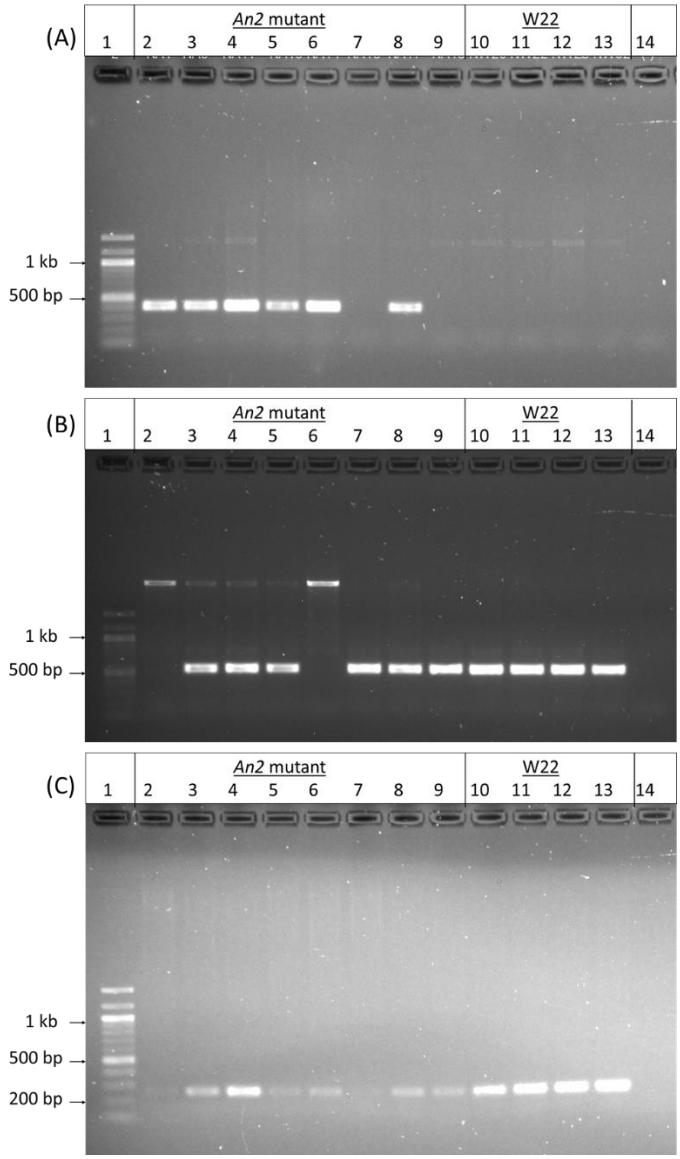
The role of kauralexins in maize defences against insect pests was further investigated to determine the reasons for the results observed for the non-choice and choice herbivory assays. Studies by Schmelz *et al.* (2011) and Dafoe *et al.* (2011), focused on the European corn borer, *O. nubilalis.* In the study by Schmelz *et al.,* non-choice and choice herbivory assays were conducted to determine whether kauralexins acted as antifeedants and / or inhibited insect growth. The non-choice herbivory assays comprised of the larvae feeding on artificial medium which had a range of different concentrations of the kauralexins A3 and B3. The

choice herbivory assays comprised of larvae feeding on stems which were untreated or treated with 50 µg·g^{-1} fresh weight of kauralexins A3 and B3. Kauralexins were found to not significantly inhibit insect growth in the non-choice assays but did significantly influence feeding preference in the choice assay (Schmelz *et al.* 2011). It is important to note that in neither the non-choice or choice feeding assays were the kauralexins synthesized *in planta* but were artificially applied or included, which means that total kauralexins applied may have been significantly more than what the plants would naturally produce. This is in contrast with the experiments conducted in the current study in which kauralexins were produced *in planta* (Section 4.3.1 and Section 4.3.2).

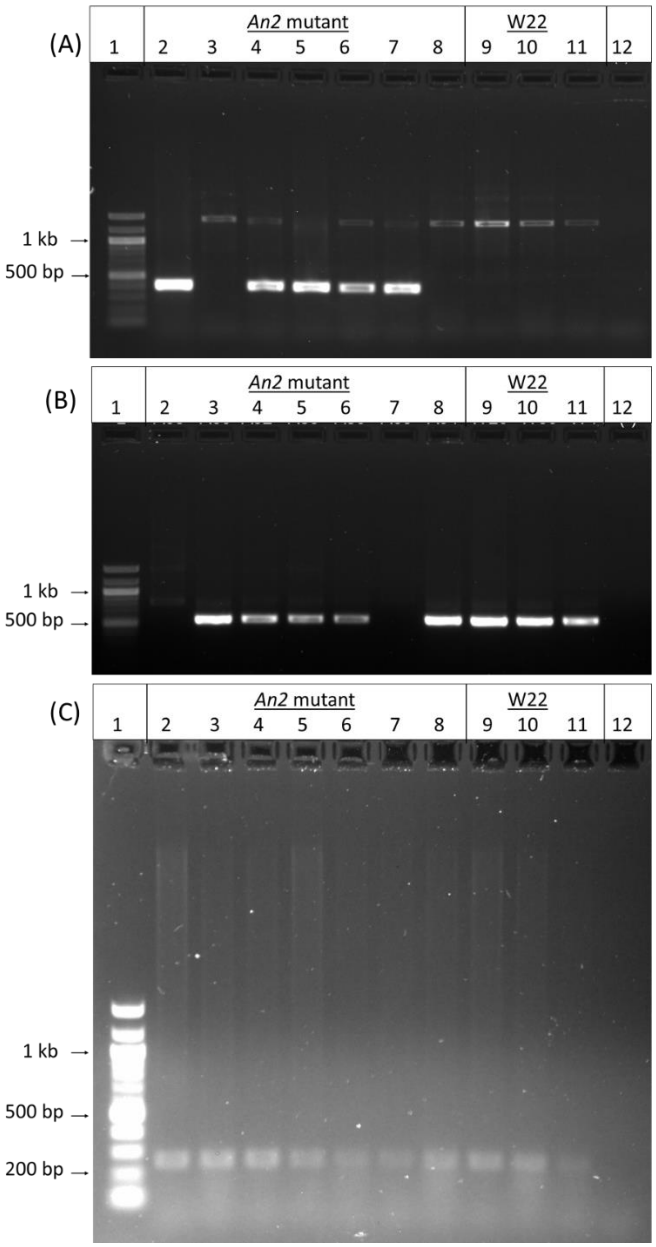
In the study by Dafoe *et al.* (2011), the kauralexin induced defence response was quantified and non-choice herbivory assays were conducted in order to determine whether larval growth and feeding was significantly influenced by kauralexins produced *in planta*. Total kauralexins were shown to significantly increase over a 24 h period with larval feeding. This confirmed that the larvae induced a defence response in the maize plants, which included the synthesis of kauralexins. The total kauralexin levels in damaged stem tissue was only 0.2 µg·g^{-1} after the 24 h period, which is much less than the 50 µg·g^{-1} total kauralexins applied in the Schmelz *et al.* study. Even though kauralexin synthesis was upregulated by larval feeding, the growth of larvae was unaffected which is the same result as which was observed in the current study. The authors speculated that this may be due to mechanisms the European corn borer possesses, which enables it to overcome these defences (Dafoe *et al.* 2011). In the current study the fall armyworm, also a lepidopteran pest like the European corn borer, is used to investigated the role of kauralexins in maize defences, similar results to the Dafoe *et al.* (2011) were obtained. The levels of kauralexins induced *in planta* by larval feeding may be too low to have an effect on the larvae over these short periods of time and the larvae used may be equipped with certain detoxification mechanisms that enable them to overcome these plant defences.

The fall armyworm is a polyphagous pest, making use of a large variety of plant hosts which differ in defence compounds. This insect pest is divided into two strains, the maize and the rice strains, which differs genetically from one another and in their responses to pesticides and host preference (Pashley *et al.* 1985, Pashley 1986, Pashley *et al.* 1987, Pashley 1988). When comparing the nuclear genomes, significant differences in detoxification and digestion genes have been found between the strains (Gouin *et al.* 2017). In the current study the strains were mixed (refer to Chapter 2 of this study). Although signs of hybrid ancestry are present, the fall armyworm used in the current study is most likely to comprise of majority maize strain genes. This is because there are no typical rice strain hosts affected in Africa, only maize strain hosts are affected and along with this the presence of pure rice strain is very rare (Nagoshi 2019). The significance of this is that the detoxification and digestion genes related to the maize strain are most likely present and in the majority. The maize strain along with its specialised cassette of detoxification and digestion genes may be the reason why the presence or absence of kauralexins did not make a difference to the feeding or mass gain of larvae. Kauralexins are naturally synthesized, maize secondary metabolites and thus would be a class of compounds common to the fall armyworm to need to metabolize and overcome. Kauralexins therefore do not play a significant role in maize defences against fall armyworm second and third instar larvae.

4.5 Supplementary Information



Supplementary Figure 4.1: Maize *An2* screening of maize samples collected from the non-choice herbivory assays. (A) Mutant *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and JGP3 primers). (B) Functional *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and 1815R primers). (C) Positive control PCR using banana actin primers of mutant *An2* and W22 maize (B.actin F and B.actin R primers). Lanes 2 – 9: An2 mutant maize; Lanes 10 – 13: W22 maize; Lane 14: non-template/water control; Lane 1: 100 bp DNA Ladder (New England Biolabs). PCR products were resolved on a 1% TAE agarose gel subjected to 80V for 50 min with ethidium bromide.



Supplementary Figure 4.2: Maize *An2* screening of maize samples collected from choice herbivory assay 1. (A) Mutant *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and JGP3 primers). (B) Functional *An2* gene PCR amplification of mutant *An2* and W22 maize (F312 and 1815R primers). (C) Positive control PCR using banana actin primers of mutant *An2* and W22 maize (B.actin F and B.actin R primers). Lanes 2 – 8: An2 mutant maize; Lanes 9 – 11: W22 maize; Lane 12: non-template/water control; Lane 1: 100 bp DNA Ladder (New England Biolabs). PCR products were resolved on a 1% TAE agarose gel subjected to 80V for 50 min with ethidium bromide.

4.6 References

Bebber, D. P., T. Holmes, D. Smith, and S. J. Gurr. 2014. Economic and physical determinants of the global distributions of crop pests and pathogens. New Phytologist 202: 901-910.

Block, A. K., M. M. Vaughan, E. A. Schmelz, and S. A. Christensen. 2019. Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). Planta 249: 21-30.

Botha, A. S., A. Erasmus, H. du Plessis, and J. Van den Berg. 2019. Efficacy of Bt maize for control of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in South Africa. Journal of Economic Entomology 112: 1260-1266.

Dafoe, N. J., A. Huffaker, M. M. Vaughan, A. J. Duehl, P. E. Teal, and E. A. Schmelz. 2011. Rapidly induced chemical defenses in maize stems and their effects on short-term growth of *Ostrinia nubilalis*. Journal of Chemical Ecology 37: 984-991.

Day, R., P. Abrahams, M. Bateman, T. Beale, V. Clottey, M. Cock*, et al.* 2017. Fall armyworm: impacts and implications for Africa. Outlooks on Pest Management 28: 196-201.

Erickson, J. M. 1973. The utilization of various *Asclepias* species by larvae of the monarch butterfly, *Danaus plexippus*. Psyche 80: 230-244.

Fox, J. 2005. Getting started with the R commander: a basic-statistics graphical user interface to R. Journal of Statistical Software 14: 1-42.

Goergen, G., P. L. Kumar, S. B. Sankung, A. Togola, and M. Tamo. 2016. First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae), a new alien invasive pest in West and Central Africa. PloS One 11: e0165632. https://doi.org/10.1371/journal.pone.0165632

Gouin, A., A. Bretaudeau, K. Nam, S. Gimenez, J. M. Aury, B. Duvic*, et al.* 2017. Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. Scientific Reports 7: 11816. https://doi.org/10.1038/s41598-017-10461-4

Harris, L. J., A. Saparno, A. Johnston, S. Prisic, M. Xu, S. Allard*, et al.* 2005. The maize *An2* gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. Plant Molecular Biology 59: 881-894.

Huang, F., J. A. Qureshi, R. L. Meagher Jr, D. D. Reisig, G. P. Head, D. A. Andow*, et al.* 2014. Cry1F resistance in fall armyworm *Spodoptera frugiperda*: single gene versus pyramided Bt maize. PloS One 9: e112958. https://doi.org/10.1371/journal.pone.0112958

Jacobs, A., A. V. Vuuren, and I. H. Rong. 2018. Characterisation of the fall armyworm (*Spodoptera frugiperda* J. E. Smith) (Lepidoptera: Noctuidae) from South Africa. African Entomology 26: 45-49.

Luginbill, P. 1928. The fall army worm, vol. 34, US Department of Agriculture.

Meyer, J., D. K. Berger, S. A. Christensen, and S. L. Murray. 2017. RNA-Seq analysis of resistant and susceptible sub-tropical maize lines reveals a role for kauralexins in resistance to grey leaf spot disease, caused by *Cercospora zeina*. BMC Plant Biology 17: 1-20. https://doi.org/10.1186/s12870-017-1137-9

Nagoshi, R. N. 2019. Evidence that a major subpopulation of fall armyworm found in the Western Hemisphere is rare or absent in Africa, which may limit the range of crops at risk of infestation. PloS One 14: e0208966. https://doi.org/10.1371/journal.pone.0208966

Nagoshi, R. N., I. Dhanani, R. Asokan, H. Mahadevaswamy, C. M. Kalleshwaraswamy, and R. L. Meagher. 2019. Genetic characterization of fall armyworm infesting South Africa and India indicate recent introduction from a common source population. PloS One 14: e0217755. https://doi.org/10.1371/journal.pone.0217755

Nagoshi, R. N., D. Koffi, K. Agboka, K. A. Tounou, R. Banerjee, J. L. Jurat-Fuentes*, et al.* 2017. Comparative molecular analyses of invasive fall armyworm in Togo reveal strong similarities to populations from the eastern United States and the Greater Antilles. PLoS One 12: e0181982. https://doi.org/10.1371/journal.pone.0181982

Oerke, E.-C. 2006. Crop losses to pests. The Journal of Agricultural Science 144: 31-43.

Pashley, D. P. 1986. Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): a sibling species complex? Annals of the Entomological Society of America 79: 898-904.

Pashley, D. P. 1988. Quantitative genetics, development, and physiological adaptation in host strains of fall armyworm. Evolution 42: 93-102.

Pashley, D. P., S. J. Johnson, and A. N. Sparks. 1985. Genetic population structure of migratory moths: the fall armyworm (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 78: 756-762.

Pashley, D. P., S. S. Quisenberry, and T. Jamjanya. 1987. Impact of fall armyworm (Lepidoptera: Noctuidae) host strains on the evaluation of bermuda grass resistance. Journal of Economic Entomology 80: 1127-1130.

Schmelz, E. A., F. Kaplan, A. Huffaker, N. J. Dafoe, M. M. Vaughan, X. Ni*, et al.* 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. Proceedings of the National Academy of Sciences 108: 5455-5460.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9: 671.

Shiferaw, B., B. M. Prasanna, J. Hellin, and M. Bänziger. 2011. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Security 3: 307-327.

Stokstad, E. 2017. New crop pest takes Africa at lightning speed, pp. 473-474, Science. American Association for the Advancement of Science.

Tholl, D. 2015. Biosynthesis and biological functions of terpenoids in plants. Biotechnology of Isoprenoids: 148:63-106. DOI: 10.1007/10_2014_295.

Valadez-Lira, J. A., J. M. Alcocer-Gonzalez, G. Damas, G. Nunez-Mejia, B. Oppert, C. Rodriguez-Padilla*, et al.* 2012. Comparative evaluation of phenoloxidase activity in different larval stages of four lepidopteran pests after exposure to *thuringiensis Bacillus*. Journal of Insect Science 12: 80.

Van den Berg, N., B. G. Crampton, I. Hein, P. R. Birch, and D. K. Berger. 2004. High-throughput screening of suppression subtractive hybridization cDNA libraries using DNA microarray analysis. Biotechniques 37: 818-824.

Vaughan, M. M., S. Christensen, E. A. Schmelz, A. Huffaker, H. J. Mcauslane, H. T. Alborn*, et al.* 2015. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. Plant, Cell & Environment 38: 2195-2207.

Veenstra, A., N. Moola, S. Wighard, J. Korsman, S. A. Christensen, M. S. Rafudeen*, et al.* 2019. Kauralexins and zealexins accumulate in sub-tropical maize lines and play a role in seedling resistance to *Fusarium verticillioides*. European Journal of Plant Pathology 153: 223-237.

Wighard, S. 2017. Characterisation of a maize mutant deficient in antifungal kauralexin accumulation. University of Cape Town.

Wu, Z., T. Wang, W. Huang, and Y. Qu. 2001. A simplified method for chromosome DNA preparation from filamentous fungi. Mycosystema 20: 575.

Yu, S. 1983. Age variation in insecticide susceptibility and detoxification capacity of fall armyworm (Lepidoptera: Noctuidae) larva. Journal of Economic Entomology 76: 219-222.

Zalucki, M. P., and L. P. Brower. 1992. Survival of first instar larvae of *Danaus plexippus* (Lepidoptera: Danainae) in relation to cardiac glycoside and latex content of *Asclepias humistrata* (Asclepiadaceae). Chemoecology 3: 81-93.

Zalucki, M. P., L. P. Brower, and A. Alonso-M. 2001. Detrimental effects of latex and cardiac glycosides on survival and growth of first-instar monarch butterfly larvae *Danaus plexippus* feeding on the sandhill milkweed *Asclepias humistrata*. Ecological Entomology 26: 212-224.

Zalucki, M. P., A. R. Clarke, and S. B. Malcolm. 2002. Ecology and behavior of first instar larval Lepidoptera. Annual Review of Entomology 47: 361-393.

Chapter 5: Conclusion

5.1 Introduction

Maize is an important crop plant in Africa as it serves as the staple food crop for over 200 million people on the continent (Day *et al.* 2017). Pests and pathogens endanger the stability of maize production and the food security offered by this crop (Oerke 2006). One such pest is the fall armyworm (*Spodoptera frugiperda*) which has recently been introduced to Africa (2016) and South Africa (2017) (Goergen *et al.* 2016, Jacobs *et al.* 2018). This lepidopteran pest can be divided into two strains, the rice and maize strains, which most importantly differ in plant host preference, resistance to chemical plant defences and pesticides (Pashley and Martin 1987, Pashley *et al.* 1987, Nagoshi and Meagher 2004). The recent introduction of this pest means that current control strategies available in Africa may not be effective in controlling fall armyworm outbreaks (Day *et al.* 2017). Investigation into maize defence compounds that are synthesized in response to fall armyworm feeding may give clues as to which compounds to focus on in developing more resistant maize lines.

Kauralexins are maize secondary metabolites which have demonstrated antimicrobial activity against important fungal pathogens, namely various *Fusarium* species, and have been shown to be synthesized in response to insect herbivory (Harris *et al.* 2005, Schmelz *et al.* 2011, Vaughan *et al.* 2015). The enzyme *ent*-copalyl diphosphate synthase (An2) plays an integral role in the kauralexin biosynthesis pathway, with gene-knock-out experiments of this enzyme resulting in near complete loss of this pathway (Harris *et al.* 2005, Ding *et al.* 2019, Veenstra *et al.* 2019). Research into the amino acid sequence and protein structure of this enzyme across various maize inbred lines may give clues as to why differing levels of kauralexins have been observed in these maize inbred lines (Schmelz *et al.* 2011, Ding *et al.* 2019). Kauralexins may play an integral role in fall armyworm larvae resistance and with differing levels of kauralexins potentially resulting in differing levels of pest and pathogen resistance, this class of maize defence compounds is a good candidate for investigation.

5.2 The importance of determining strain identity in a laboratory reared fall armyworm culture

The fall armyworm laboratory reared cultures which were used in this study were determined to consist of rice strain, maize strain and interstrain hybrid individuals using the *COI* and *Tpi* genetic markers (Goergen *et al.* 2016, Nagoshi *et al.* 2017, Jacobs *et al.* 2018). These genetic markers when used in isolation did not allow for consistently accurate strain identification, especially when dealing with fall armyworm larvae when interstrain hybridisation is present in the specimen's ancestry and the sex of the specimen is unknown (Meagher and Nagoshi 2012). When both genetic markers are used, interstrain hybridisation and potential clues to hybridisation in the specimen's ancestry in the form of discordant markers can be identified, but ultimately this is not consistently accurate when dealing with a laboratory reared culture which may be frequently hybridising and inbreeding.

The need to develop a nuclear based genetic marker to be used in conjunction with the *COI* and *Tpi* genetic markers was highlighted as many of the larvae screened from laboratory reared cultures exhibited discordance between the *COI* and *Tpi* markers, unable to define the strain identity of these specimens. This is important because the effect of interstrain hybridisation is largely unknown (Nagoshi 2019). The detoxification genes, of which significant differences have been identified between the strains, are potentially being traded with every mating recombination event (Gouin *et al.* 2017). Inaccurate strain identification may lead to these genetic changes being unaccounted for, leading to potentially irreplicable experiments and the application of ineffective control strategies that result in a loss of time and valuable resources.

5.3 A look at An2 across different maize inbred lines

The *ent*-copalyl diphosphate synthase enzyme which is responsible for converting geranyl-geranyl diphosphate to *ent*-copalyl diphosphate, is integral in synthesizing kauralexins (Harris *et al.* 2005, Vaughan *et al.* 2015, Murphy *et al.* 2018, Ding *et al.* 2019). This enzyme was investigated across six maize inbred lines (B73, W22, CO387, CML333, Tx303, PE0075) to determine whether variation in amino acid sequence was present and whether this variation may lead to differences in binding affinity of the enzyme to the ligand, geranyl-geranyl diphosphate (GGPP). Variation in amino acid sequence and predicted binding affinity was observed. The maize inbred lines, W22 and CML333, were reported to be amongst the highest total kauralexin synthesizing maize inbred lines in previous research articles (Schmelz *et al.* 2011, Ding *et al.* 2019). This correlated with the predictive ligand docking in which CML333 and W22 An2 proteins exhibited the highest binding affinity to the ligand compared to the remaining 4 maize inbred lines evaluated (Schmelz *et al.* 2011, Ding *et al.* 2019). Investigating amino acid sequence variation and comparing protein structures of An2 enzyme for different maize inbred lines showed promising results. Evaluating protein structure of enzymes critical in plant defence synthesis can be useful in identifying maize inbred lines which synthesize higher levels of important defence compounds, such as kauralexins.

5.4 The role of kauralexins in maize defences against the fall armyworm

The role of kauralexins was investigated by using kauralexin synthesizing maize, the W22 maize inbred line, and a low to no kauralexin synthesizing maize, An2 mutant maize. Utilizing non-choice and choice herbivory assays it was determined that kauralexins do not play a significant role in maize defences against fall armyworm larval feeding *in planta*. There were no significant differences observed in the mass gain of the larvae and the total maize leaf area consumed on either maize line. This result correlated with results in previously published

literature in which it has also been shown that *in planta* kauralexins do not play a significant role in maize defences against European corn borer larvae (Dafoe *et al.* 2011, Schmelz *et al.* 2011).

5.5 Research shortcomings

The laboratory reared culture and the variation observed in terms of seemingly pure strain specimens and interstrain hybrids is complicated by the low number of specimens screened. This was initially not part of the original study, but when variation in *COI* and *Tpi* strain identity was observed in the laboratory reared culture, it was then given more attention. Screening higher numbers of fall armyworm specimens would have given a clearer picture of the hybridisation which was taking place with continuous inbreeding in the culture. Many of the specimens were also of unknown sex. This means that if these specimens were not heterozygous for the specific *Tpi* SNP marker, it was unknown whether these samples were in fact hemizygous or homozygous for the *Tpi* strain identity (Nagoshi 2010). This is in part, a pitfall of the *Tpi* marker itself, but could have partially been avoided if the specimens screened were consistently pupae. Pupae can be morphologically defined as either male or female. The detoxification genes that were amplified and shown to be variable across hybrid samples highlighted the importance of accurate strain identification and the potential effects of hybridisation. The pitfall of this part of the study was the lack of sequence data for these amplified sequences. Sequencing these detoxification genes would have confirmed the reliability of the PCR's performed. More samples screened using these detoxification markers would have also been beneficial to further confirm the shifting genetics of the hybridising fall armyworm cultures.

The evaluation An2 protein sequences and structures across multiple maize lines did have some pitfalls, because the nature of this study is predictive. This part of the study was

completely reliant on previously published research and predictive protein modelling. The ability to construct accurate maize An2 protein structures was very much constrained by the lack of a maize An2 protein template, a maize An2 protein for which the crystalline structure had been determined. The template protein used to model the maize An2 proteins after was an *ent*-copalyl diphosphate synthase protein found in *A. thaliana* (Köksal *et al.* 2011). The template protein and proteins which were to be modelled after it, were not highly similar (42% similarity), which means the resulting protein structures are potentially less accurate than one would hope for. Less accurate protein structures lead to less accurate protein-ligand interactions which would then lower the accuracy of the predicted binding affinity of the An2 protein and the GGPP ligand.

The mutant An2 maize used to investigate the role of kauralexins in maize defences against fall armyworm larvae was not consistently homozygous for the mutation. Some maize specimens were determined to be heterozygous for the mutation, while a small minority were observed to be completely lacking the mutation. Ideally the mutant An2 maize should have been homozygous for the mutation in order to make a strong comparison with the W22 maize lacking the mutation (synthesizing normal kauralexin levels). Although this is a pitfall to these experiments, the results observed are still trustworthy as there was no variation observed in the choice herbivory assay 2, which then used only homozygous An2 maize. Variation in larval size between bioassays may have been the cause of repeated experiments being significantly different from one another and not being able to have these data sets combined although the same experimental protocols were followed. This observation may instead be due to variation in the genetic structure of the fall armyworm culture used as the culture continuously inbred and hybridised over the course of the different experiments performed.

5.6 Future prospects

Fall armyworm strain identification is in need of more nuclear based genetic markers. Identifying genes that play integral roles in plant host preference and chemical resistance associated with the different strains could lead to the development of strain specific markers. Detoxification and gustatory receptor genes are promising candidates which future studies could further investigate (Gouin *et al.* 2017). Future studies could involve performing real-time quantitative PCR tests of some of these detoxification genes from fall armyworm specimens which have fed on various maize genotypes to identify which detoxification genes that are upregulated in response to maize defences. The detoxification gene profiles can be compared between the strains to identify which detoxification genes result in fitness advantages and which genes could serve as genetic markers for defining hybrid strain samples as either being more maize or rice strain leaning.

Whole genome sequencing of fall armyworm specimens is another approach, which would aid in further defining the strains. This approach will also be beneficial in understanding the pest's invasive and adaptive mechanisms, as well as identifying the genetic basis of developing insecticide resistance (Kirk *et al.* 2013, Rius *et al.* 2015, Poelchau *et al.* 2016). Currently there are nine genome assemblies available for fall armyworm (Yang *et al.* 2021). A population genomics approach can be utilised to further define the strains. Representative specimens from maize and rice strain populations from different American regions can be collected, genomes sequenced and genome wide association studies (GWAS) performed to identify single nucleotide polymorphisms (SNPs) (Korte and Farlow 2013). These SNPs would then be evaluated to identify which associate with either strain, where pure maize or rice strains may be defined by statistically categorising the occurrence of the strain specific SNPs. Defining what "pure" strains are using individual specimens is likely not possible due to high levels of heterozygosity between specimens within each strain. The aim of whole genome

sequencing in this context would be to accurately define strain identity and identify more reliable molecular markers enabling more repeatable research, accurate predictions of the effects of fall armyworms populations in new areas and further aid in developing effective chemical control strategies.

To further investigate the amino acid and protein structure variation of the different enzymes involved in kauralexin synthesis, crystalline protein structures need to be determined for the enzymes in question. High similarity between the template protein amino acid sequence and the amino acid sequences used to construct the homology models (>90%), result in reliable models and more accurate predictive results. Once reliable maize template proteins are available, the potential for uncovering the most optimal kauralexin synthesising enzymes is available. Future studies can also further investigate the findings of this study. A glutamine at position 460 on the An2 amino acid sequence was identified to be common amongst the higher kauralexin synthesizing maize inbred lines, whereas lower kauralexin synthesizing maize lines exhibited a glutamic acid amino acid base at this position. The influence of this amino acid base and its impact on kauralexin synthesis can be experimentally determined through the use of CRISPR-cas to induce a missense mutation (Liu *et al.* 2020). A missense mutation in the An2 gene of W22 maize, changing the glutamine amino acid at position 460 of the amino acid sequence to a glutamic acid. This maize can then be subject to insect herbivory and fungal pathogens, resulting in kauralexin synthesis. Performing liquid chromatography – mass spectroscopy (LC-MS) on this plant tissue would reveal the influence of the missense mutation total kauralexin synthesis and whether the amino acid base is critical in the enzyme's catalytic activity. The other enzymes part of the kauralexin synthesis pathway can be investigated in a similar way across the different maize inbred lines, identifying single amino acid bases which influence catalytic activity. Identifying unique single amino acids in these enzymes critical in kauralexin synthesis may lead to the development of new genetic markers which maize breeders could utilise to incorporate higher kauralexin synthesising maize into their breeding programs.

Although kauralexins were found to not be effective as a chemical defence *in planta* against fall armyworm larvae, kauralexins may still be useful at higher concentrations. Future studies could perform feeding assays using artificial diet containing varying concentrations of kauralexins. This approach was not used in this study due to the focus of the kauralexin synthesis pathway *in planta* and the additional costs associated with either purchasing the different kauralexin compounds or developing an expression system to artificial synthesise them. The use of kauralexins as a foliar spray at higher concentrations could be also investigated. Kauralexins when applied as a high concentration solution onto maize stalks resulted in significantly lower feeding by European corn borer larvae in previous studies (Schmelz *et al.* 2011). If a kauralexin foliar spray was found to be effective, this could be a very useful addition to integrated control strategies against this pest.

5.7 Conclusion

Maize defence biosynthesis pathways, such as the kauralexin biosynthesis pathway, are important to understand to continue developing crops resistant to various pests and pathogens in the effort to further secure our food systems. Variations in the An2 enzyme, critical in kauralexin synthesis, were identified across multiple maize inbred lines and could be correlated to total kauralexin synthesis data from previous studies. This finding means that the protein structure of the An2 enzyme is important in the total kauralexin synthesis potential in maize and deserves further research. Although kauralexins were found to play an insignificant role in maize defences against fall armyworm *in planta*, these compounds are still a promising compound class to investigate for their role in maize defences against other pests and pathogens.

5.8 References

Dafoe, N. J., A. Huffaker, M. M. Vaughan, A. J. Duehl, P. E. Teal, and E. A. Schmelz. 2011. Rapidly induced chemical defenses in maize stems and their effects on short-term growth of *Ostrinia nubilalis*. Journal of Chemical Ecology 37: 984-991.

Day, R., P. Abrahams, M. Bateman, T. Beale, V. Clottey, M. Cock*, et al.* 2017. Fall armyworm: impacts and implications for Africa. Outlooks on Pest Management 28: 196-201.

Ding, Y., K. M. Murphy, E. Poretsky, S. Mafu, B. Yang, S. N. Char*, et al.* 2019. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. Nature Plants 5: 1043-1056.

Goergen, G., P. L. Kumar, S. B. Sankung, A. Togola, and M. Tamo. 2016. First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae), a new alien invasive pest in West and Central Africa. PloS One 11: e0165632. https://doi.org/10.1371/journal.pone.0165632

Gouin, A., A. Bretaudeau, K. Nam, S. Gimenez, J. M. Aury, B. Duvic*, et al.* 2017. Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. Scientific Reports 7: 11816. https://doi.org/10.1038/s41598-017-10461-4

Harris, L. J., A. Saparno, A. Johnston, S. Prisic, M. Xu, S. Allard*, et al.* 2005. The maize *An2* gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. Plant Molecular Biology 59: 881-894.

Jacobs, A., A. V. Vuuren, and I. H. Rong. 2018. Characterisation of the fall armyworm (*Spodoptera frugiperda* J. E. Smith) (Lepidoptera: Noctuidae) from South Africa. African Entomology 26: 45-49.

Kirk, H., S. Dorn, and D. Mazzi. 2013. Molecular genetics and genomics generate new insights into invertebrate pest invasions. Evolutionary Applications 6: 842-856.

Köksal, M., H. Hu, R. M. Coates, R. J. Peters, and D. W. Christianson. 2011. Structure and mechanism of the diterpene cyclase *ent*-copalyl diphosphate synthase. Nature Chemical Biology 7: 431-433.

Korte, A., and A. Farlow. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9: 1-9. https://doi.org/10.1186/1746-4811-9-29

Liu, H.-J., L. Jian, J. Xu, Q. Zhang, M. Zhang, M. Jin*, et al.* 2020. High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. The Plant Cell 32: 1397-1413.

Meagher, R. L., and R. N. Nagoshi. 2012. Differential feeding of fall armyworm (Lepidoptera: Noctuidae) host strains on meridic and natural diets. Annals of the Entomological Society of America 105: 462-470.

Murphy, K. M., L.-T. Ma, Y. Ding, E. A. Schmelz, and P. Zerbe. 2018. Functional characterization of two class II diterpene synthases indicates additional specialized diterpenoid pathways in maize (*Zea mays*). Frontiers in Plant Science 9: 1542. https://doi.org/10.3389/fpls.2018.01542

Nagoshi, R. N. 2010. The fall armyworm triose phosphate isomerase (*Tpi*) gene as a marker of strain identity and interstrain mating. Annals of the Entomological Society of America 103: 283-292.

Nagoshi, R. N. 2019. Evidence that a major subpopulation of fall armyworm found in the Western Hemisphere is rare or absent in Africa, which may limit the range of crops at risk of infestation. PloS One 14: e0208966. https://doi.org/10.1371/journal.pone.0208966

Nagoshi, R. N., and R. L. Meagher. 2004. Seasonal distribution of fall armyworm (Lepidoptera: Noctuidae) host strains in agricultural and turf grass habitats. Environmental Entomology 33: 881-889.

Nagoshi, R. N., D. Koffi, K. Agboka, K. A. Tounou, R. Banerjee, J. L. Jurat-Fuentes*, et al.* 2017. Comparative molecular analyses of invasive fall armyworm in Togo reveal strong similarities to populations from the eastern United States and the Greater Antilles. PLoS One 12: e0181982. https://doi.org/10.1371/journal.pone.0181982

Oerke, E.-C. 2006. Crop losses to pests. The Journal of Agricultural Science 144: 31-43.

Pashley, D. P., and J. A. Martin. 1987. Reproductive incompatibility between host strains of the fall armyworm (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 80: 731-733.

Pashley, D. P., S. S. Quisenberry, and T. Jamjanya. 1987. Impact of fall armyworm (Lepidoptera: Noctuidae) host strains on the evaluation of bermuda grass resistance. Journal of Economic Entomology 80: 1127-1130.

Poelchau, M. F., B. S. Coates, C. P. Childers, A. A. P. de Leon, J. D. Evans, K. Hackett, *et al.* 2016. Agricultural applications of insect ecological genomics. Current Opinion in Insect Science 13: 61-69.

Rius, M., S. Bourne, H. G. Hornsby, and M. A. Chapman. 2015. Applications of next-generation sequencing to the study of biological invasions. Current Zoology 61: 488-504.

Schmelz, E. A., F. Kaplan, A. Huffaker, N. J. Dafoe, M. M. Vaughan, X. Ni, *et al.* 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. Proceedings of the National Academy of Sciences 108: 5455-5460.

Vaughan, M. M., S. Christensen, E. A. Schmelz, A. Huffaker, H. J. Mcauslane, H. T. Alborn*, et al.* 2015. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. Plant, Cell & Environment 38: 2195-2207.

Veenstra, A., N. Moola, S. Wighard, J. Korsman, S. A. Christensen, M. S. Rafudeen*, et al.* 2019. Kauralexins and zealexins accumulate in sub-tropical maize lines and play a role in seedling resistance to *Fusarium verticillioides*. European Journal of Plant Pathology 153: 223-237.

Yang, P., D. Wang, W. Guo, and L. Kang. 2021. FAWMine: An integrated database and analysis platform for fall armyworm genomics. Insect Science 28: 590-601.

Summary

Kauralexins are a class of terpenoid, secondary metabolites synthesized in maize. These compounds have exhibited antimicrobial and insect antifeedant activities. The fall armyworm, *Spodoptera frugiperda,* is a notorious maize insect pest which has the potential to negatively impact maize production. The purpose of this study is to investigate the kauralexin biosynthetic pathway, variation in kauralexin synthesis across maize inbred lines and the role of this pathway in maize defences against fall armyworm larvae. Molecular markers used to determine fall armyworm strain identity were also investigated and the reliability of the commonly used markers discussed.

The fall armyworm is comprised of two strains, the maize and rice strain. These strains are indistinguishable morphologically and therefore, molecular markers are relied upon to distinguish these strains. The widely used molecular markers, *cytochrome oxidase I* and *triosephosphate isomerase* markers, are discussed in this study. The uses and pitfalls of these markers, the need for accurate strain identification and implications for strain hybridisation in a laboratory reared culture in a research context is discussed. Laboratory reared fall armyworm specimens were screened using these markers and PCR amplification of detoxification genes unique to either strain performed. The results showed that using either marker in isolation could result in misidentification of interstrain hybrids and highlighted the need for new nuclear based markers.

Kauralexin synthesis differs across multiple maize inbred lines. The enzyme, Anther ear 2 (An2), plays an integral role in kauralexin synthesis, with previous studies showing a knock-out of this gene resulting in near complete reduction of kauralexin levels. Variation in this enzyme was investigated as the potential reason for the varying kauralexin levels observed in the different maize inbred lines. Six maize inbred line were investigated by comparing the

amino acid sequences of the An2 proteins, performing protein modelling and ligand docking. The ligand docking revealed variation in binding affinity of the ligand to the An2 proteins which was then linked to the variation observed in total kauralexin synthesis. Total kauralexin levels correlated with the differing binding affinities observed.

The role of kauralexins in maize defences against fall armyworm larvae was investigated through the use of choice and non-choice herbivory assays. These herbivory assays compared larval feeding on maize which synthesized kauralexins and maize which synthesized low to no kauralexins (*An2* gene knock-out). The parameters measured in these experiments were leaf area consumed and larval mass gain. Kauralexins did not play a significant role in maize defences *in planta* against fall armyworm larvae. These compounds did not significantly influence larval feeding or mass gain and preference of the larvae was unaffected.