

THE GENETIC CONTROL OF AGRICULTURAL PESTS

(PLUTELLA XYLOSTELLA, L. AND TRIBOLIUM CASTANEUM, HERBST)

Matthew William Gregory

LADY MARGARET HALL



DEPARTMENT OF ZOOLOGY

UNIVERSITY OF OXFORD

A Thesis submitted in fulfilment

of the Degree of

Doctor of Philosophy

TRINITY TERM 2015

ACKNOWLEDGEMENTS

Thanks to everyone at Oxitec for an enjoyable experience (I intend to thank you in person!).

Many of the Chapters in this thesis would not have been possible without the collaboration of other scientists; sharing their time, equipment and wisdom. Special thanks go to Helen White-Cooper for her advice throughout and her help with the RISH at Cardiff. Thanks go to the Brown lab at the University of Kansas for providing me with *Tribolium* and rearing advice. Thanks to Nina Alphey and Mike Bonsall for organizing a memorable experience — at the Royal Society Summer Science Exhibition — providing the opportunity to interact with the public and discuss our research and its implications.

I would also like to thank my funding council for providing exceptional training opportunities throughout my DPhil; particularly the SysMIC course which provided advanced training in maths and computing based around intellectually stimulating biological examples.

A final thanks to my supervisors Seb Shimeld, Neil Morrison and Luke Alphey.

DECLARATION OF AUTHENTICITY

This work was made possible through the financial support of the Biotechnology and Biological Sciences Research Council (BBSRC) CASE studentship scheme and Oxitec Ltd. All work presented in this thesis was conducted by the author, except where explicitly noted in the text. None of the following has been submitted previously for a qualification at the University of Oxford, or any other institute of higher education.

ABSTRACT

The interaction between an increasing global human population and a number of important long-term trends and issues are putting strain on food production. Insects represent a significant food security concern causing up to 15% of global crop losses. Conventional chemical methods are ineffective; inducing resistance and degrading the environment. Sustainable alternatives are sought.

The sterile insect technique provides a sustainable solution. Genetic engineering can augment this historic technique by replacing radiation-induced sterilization with sperm-specific nuclease expression to introduce double-stranded DNA breaks in the gametes of mass-reared and released males. This paternal-effect system is dependent on elucidation of appropriate sperm-specific promoters and suitable chaperone-nuclease combinations. This thesis develops this technology in the SIT neglected insect orders, the Lepidoptera and Coleoptera. Specifically the diamondback moth (*Plutella xylostella*, L.) and the red flour beetle (*Tribolium castaneum*, Herbst).

I provide the foundations for a paternal-effect genetic-control-system in both species by developing a conditional sperm-specific expression system in diamondback moth and a female-specific expression system in *Tribolium*.

Mass-rearing insects for the genetic control of a species can be augmented by recent developments in RNAi. I show that the sex ratio of *Tribolium* can be adjusted by treatment with dsRNA *transformer*, producing pseudomales as an additional bonus.

In addition, an exploratory data analysis of producing transgenic lines in insects using *piggyBac* was undertaken. As well as providing a comprehensive compendium and assessment of the transgenic literature, something not yet published elsewhere, a predictive model was produced that could be very useful to a wide diversity of researchers in insect molecular biology, developmental biology, disease biology and genetics.

It is hoped that this work will contribute towards the effective control of the diamondback moth in the near term, and form a model for the sustainable control of other lepidopteran and coleopteran species through genetic pest management.

TABLE OF CONTENTS

Chapter 1 General Introduction

| | |
|---|----|
| 1.1 Background | 1 |
| 1.1.1 The problem | 1 |
| 1.1.2 Possible solutions – an integrated approach to agriculture | 3 |
| 1.1.3 The sterile insect technique | 4 |
| 1.1.4 SIT by genetic modification | 5 |
| 1.1.5 Target species | 9 |
| 1.1.6 Limitations of the traditional SIT approach | 11 |
| 1.1.7 Release of insects carrying a dominant lethal..... | 12 |
| 1.1.8 Sex sorting | 15 |
| 1.1.9 Conventional RIDL versus paternal effect | 16 |
| 1.1.10 Mathematical models for the application of sterile insects | 23 |
| 1.1.11 Construct design..... | 25 |
| 1.2 Aim..... | 28 |
| 1.3 Objectives..... | 29 |

Chapter 2 Insect transformation with *piggyBac*: getting the number of injections just right

| | |
|---|----|
| 2.1. Introduction..... | 32 |
| 2.1.1 Insect transformation using <i>piggyBac</i> | 32 |
| 2.2. Results and Discussion..... | 37 |

| | |
|---|----|
| 2.2.1 Systematic review of transformation efficiencies in published literature | 37 |
| 2.2.2 Transformation efficiency by insect order and species from published data | 38 |
| 2.2.3 Analysis of an extensive unpublished dataset of <i>piggyBac</i> transgenesis experiments | 42 |
| 2.2.4 Goldilocks Decision-Making: how to get the number of injections just right with <i>piggyBac</i> | 52 |
| 2.3. Conclusion | 57 |
| <u>Chapter 3 What components of a microinjection mix are required for transient expression of a marker in an embryo?</u> | |
| 3.1 Introduction..... | 59 |
| 3.2 Exploratory data analysis | 61 |
| 3.3 What is required for transient expression? | 64 |
| 3.4 Power analysis and blinding | 65 |
| 3.5 Method | 66 |
| 3.6 Results | 67 |
| 3.6.1 Differences between treatments..... | 67 |
| 3.6.2 Differences in embryo transience status between consecutive days | 72 |
| 3.6.3 When should I check for embryo transience? | 75 |
| 3.6.4 Pupal transience | 77 |
| 3.7 Conclusions | 79 |
| <u>Chapter 4 Chapter 4:</u> | |
| 4.1. Introduction..... | 82 |
| 4.1.1 An overview of the paternal effect system | 82 |
| 4.1.2 Sperm development in the Lepidoptera | 84 |

| | |
|---|-----|
| 4.1.3 Sperm-specific regulatory sequences..... | 87 |
| 4.1.4 The ideal gene for the job | 88 |
| 4.1.5 Idiosyncrasies of sperm in the Lepidoptera..... | 95 |
| 4.1.6 Restriction enzymes | 95 |
| 4.1.7 Chaperone..... | 96 |
| 4.2 Hypotheses..... | 97 |
| 4.3 Results and Discussion..... | 98 |
| 4.3.1 Determination of gene expression patterns of paternal effect candidate genes in diamondback moth testes using <i>in situ</i> hybridisation..... | 98 |
| 4.3.2 <i>In situ</i> hybridisation..... | 105 |
| 4.3.3 Toward a paternal effect system in diamondback moth..... | 115 |
| 4.3.4 OX4804 – Testing 5' UTR <i>Dmprot</i> interference?..... | 133 |
| 4.3.5 OX5067 – bigger is better?..... | 142 |
| 4.3.6 OX5116 – is intronic DNA necessary for efficient transcription? | 151 |
| 4.3.7 Creating a construct with the $\beta 2$ - <i>tubulin</i> and DsRed2 transgene separated by ubiquitin | 160 |
| 4.3.8 OX5133 and OX5135 | 161 |
| 4.3.9 Diamondback moth $\beta 2$ - <i>tubulin</i> axenome and 3' UTR..... | 171 |
| 4.3.10 Five to three - our transgene in OX5196 now includes most of $\beta 2$ - <i>tubulin</i> | 177 |
| 4.4 Conclusions | 189 |
| <u>Chapter 5 Chapter 5 – The development of a repressible genetic sexing mechanism in <i>Tribolium castaneum</i></u> | |
| 5.1 Background | 195 |

| | |
|--|-----|
| 5.1.1 Suitability for SIT | 195 |
| 5.1.2 Suitability for genetic control and a sexing strain using RIDL | 197 |
| 5.1.3 Sex determination..... | 197 |
| 5.1.4 Promoter choice | 199 |
| 5.2. Objectives | 200 |
| 5.3 Results and Discussion..... | 201 |
| 5.3.1 Transgenesis in <i>Tribolium</i> | 202 |
| 5.3.2 Rearing <i>Tribolium</i> on a tetracycline-dosed diet..... | 215 |
| 5.3.2.5 Tetracycline delivery to male gonads..... | 218 |
| 5.3.3 Design of a genetic sexing construct (female-lethal) – <i>doublesex</i> or <i>transformer</i> ?..... | 221 |
| 5.3.4 OX4799 – Transgene-based, female-specific lethality-based system for genetic sexing in <i>Tribolium</i> | 224 |
| 5.3.5 OX5118, green transformants and red females expected..... | 228 |
| 5.3.6 OX5145 - What OX5118 was intended to be! | 242 |
| 5.4 Conclusions | 259 |
| 5.5 Future direction..... | 260 |
| <u>Chapter 6 Chapter 6: Insect Tra-genderesis: silencing femaleness as part of a mass-rearing programme</u> | |
| 6.1. Introduction..... | 262 |
| 6.1.1 RNAi in insects..... | 262 |
| 6.1.2 Hypotheses..... | 264 |
| 6.2 RNAi in <i>Tribolium</i> | 264 |
| 6.2.1 <i>Tribolium</i> as a model for RNAi | 264 |

| | |
|---|-----|
| 6.2.2 Experimental design considerations for <i>Tctra</i> RNAi..... | 265 |
| 6.2.3 Experimental treatment and control considerations..... | 266 |
| 6.2.4 Reproducing and building on research | 267 |
| 6.3 Conclusions | 290 |
| <u>Chapter 7 Thesis synthesis</u> | |
| 7.1 General discussion and future considerations | 294 |
| 7.2 Conclusion | 301 |
| <u>Chapter 8 Methods and Materials and Troubleshooting</u> | |
| 8.1 Insect rearing | 303 |
| 8.1.1 Diamondback moth..... | 303 |
| 8.1.2 Preparation of sugar water; tetracycline / non-tetracycline | 303 |
| 8.1.3 Larval diet..... | 303 |
| 8.1.4 <i>Tribolium castaneum</i> | 304 |
| 8.1.5 <i>Tribolium castaneum</i> rearing on a tetracycline-dosed diet | 305 |
| 8.1.6 Testis anatomy and development in diamondback moth..... | 305 |
| 8.2. Molecular Methods..... | 305 |
| 8.2.1 Preparation of dNTPs | 305 |
| 8.2.2 Extraction of genomic DNA..... | 305 |
| 8.2.3 Extracting RNA | 306 |
| 8.2.4 Reverse transcriptase PCR using RevertAid | 307 |
| 8.2.5 Quantification of DNA and RNA using GeneQant 2 | 308 |
| 8.2.6 Polymerase chain reaction (PCR) | 308 |
| 8.2.7 Primer choice | 309 |

| | |
|---|-----|
| 8.2.8 Sequencing | 309 |
| 8.2.9 Sequence alignment..... | 309 |
| 8.2.10 Agarose gel electrophoresis | 309 |
| 8.2.12 Transformation using XL-10 gold Ultracompetent cells..... | 309 |
| 8.2.13 Picking colonies for PCR screening and preparing cultures for minipreps | 310 |
| 8.2.14 Isolation of plasmid DNA from E. coli to produce a miniprep | 310 |
| 8.2.15 5'RACE using SMARTer RACE cDNA amplification kit..... | 310 |
| 8.2.16 Qiagen MinElute PCR reaction clean up | 311 |
| 8.2.17 Purification and extraction of DNA using an agarose gel | 311 |
| 8.2.18 DNA Purification using Qiagen QIAfilter Endofree Plasmid Maxi Kit | 311 |
| 8.2.19 High fidelity PCR using Q5 Polymerase (M0491)..... | 311 |
| 8.2.20 One-step RT-PCR using SuperScript III with Platinum Taq | 311 |
| 8.2.21 <i>In vitro</i> transcription of RNA using MEGAshortscript kit (T7 promoter) | 312 |
| 8.2.22 Western Blot of DsRed..... | 312 |
| 8.2.23 Leg PCR for genotyping..... | 312 |
| 8.2.24 Multiple PCR in <i>Tribolium</i> for sexing gDNA..... | 313 |
| 8.2.25 RNA <i>in situ</i> hybridisation in diamondback moth..... | 314 |
| 8.3 Generation of germ-line transgenic strains..... | 325 |
| 8.3.1 Preparation of Transformation medium or injection buffer..... | 325 |
| 8.3.2 Preparation of plasmids for microinjection | 325 |
| 8.3.3 Preparation of Injection Capillaries | 325 |
| 8.3.4 Collection and preparation of embryos for injection | 326 |
| 8.3.5 Rearing injected embryos..... | 326 |

| | |
|--|-----|
| 8.3.6 Crossing G ₀ and screening G ₁ | 327 |
| 8.3.7 Establishing stock populations | 328 |
| 8.3.8 Injections | 329 |

Chapter 9 Appendices

| | |
|---|-----|
| 9.1 Constructs | 330 |
| 9.1.1 Construct design general components and considerations | 330 |
| 9.1.2 OX4703 | 331 |
| 9.1.3 OX4700 | 333 |
| 9.1.4 OX4804 | 335 |
| 9.1.5 OX4799 | 335 |
| 9.1.6 OX5067 | 337 |
| 9.1.7 OX5116 | 338 |
| 9.1.8 OX5118 | 339 |
| 9.1.9 OX5133 | 341 |
| 9.1.10 OX5135 | 343 |
| 9.1.11 OX5145 | 343 |
| 9.1.12 OX5196 | 344 |
| 9.1.13 OX5202 | 345 |
| 9.1.14 OX4026 | 347 |
| 9.2 5' RACE of 5'UTR diamondback moth β2 tubulin | 348 |
| 9.3 Primer sequences..... | 350 |
| 9.4 Transformation efficiency using <i>piggyBac</i> literature review | 355 |
| 9.5 Protein assay for DsRed in OX5133..... | 359 |

| | |
|--|------------|
| 9.5.1 A Western blot with a monoclonal primary antibody for DsRed..... | 359 |
| 9.5.2 A Western blot with a polyclonal primary antibody for DsRed..... | 363 |
| 9.6 Optimising <i>Tribolium castaneum</i> micro-injections for insect transgenesis | 364 |
| 9.6.1 Issues for mass-rearing and adequate embryo production for injection | 364 |
| 9.6.2 Embryo mass production troubleshooting..... | 373 |
| 9.6.3 Embryo collection..... | 374 |
| 9.6.4 Embryo care..... | 374 |
| 9.6.5 Post-injection survival of embryos | 376 |
| 9.6.6 Post-sieve embryo treatment prior to injection | 378 |
| 9.7 Supplementary Methods for Chapter 2 | 383 |
| 9.7.1 Literature Review | 383 |
| 9.7.1.2 Statistical description of the transformation efficiency data | 390 |
| 9.7.1.3 Bayesian analysis of transformation efficiency – a toy example with <i>Tribolium castaneum</i> | 394 |
| 9.7.1.4 Oxitec data | 396 |
| 9.8 GLMM details..... | 403 |
| 9.9 Papers | 404 |
| 9.10 Chapter cover page photos..... | 404 |
| <u>Chapter 10 References</u> | 403 |

CHAPTER 1: GENERAL INTRODUCTION



1.1 BACKGROUND

1.1.1 THE PROBLEM

More than 200 years ago, Malthus famously asserted that the growth of the human population is geometric while the rate of change of global food production is arithmetic (Malthus, 1798). It is predicted that today's global population will continue to grow by an estimated 3 billion people over the next 40 years (Gerland *et al.*, 2014). Will global food production and distribution be able to meet this demand, avoiding a Malthusian crunch?

The interaction between an increasing population and a number of important long-term trends and issues are putting strain on food production (Beddington *et al.*, 2010; BeVier, 2012; Dube *et al.*, 2012). The scientific method provides optimism and a mechanism for producing more food, more efficiently and sustainably from a given unit area using knowledge intensive methods (Sayer & Cassman, 2013). A pertinent research area to this end is focused on mitigating pest damage to agro-ecosystems, with estimated crop losses of 15% inflicted by insects alone annually (Maxmen, 2013).

Reducing crop losses to agricultural insect pests will play a role in meeting the increase in demand and mitigating food security issues. Sustainable intensification is desired, whereby more food can be produced from an area of land while reducing the environmental impact (Godfray *et al.*, 2010; Tscharntke *et al.*, 2012). Furthermore, evidence suggests that the Borlaug effect, whereby intensification of agriculture reduces the need for extensification, is real, averting net habitat loss and environmental damage (Stevenson *et al.*, 2012; Haddad *et al.*, 2015).

This is not the first time humanity has confronted this problem. The era of the industrial revolution and post-war years experienced exponential population growth and the associated risk of undernourishment. Yet food production kept pace due to improvements in agricultural practices and commensal yield increases. This Green Revolution was heavily dependent on the flourishing chemical and petro-chemical industry (Matthews, 2006; Pingali, 2012).

Despite the success of the Green Revolution in feeding people, there were significant costs and problems (Bazuin *et al.*, 2011) often overlooked (Tscharntke *et al.*, 2012). Public concern was intensified by the publication of Silent Spring (Carson, 1962), and the problem of feeding the world with minimal environmental impact persists (Green *et al.*, 2005), like dichlorodiphenyltrichloroethane (DDT) bioaccumulation.

Protection of field crops and stored agricultural produce has historically relied on the use of chemical pest control methods (Matthews, 2006; McGraw & O'Neil, 2013). Contemporary

legislation, particularly in the European Union (EU) (Banasiaik, *et al.*, 2010; EU directive 2009/128/EC described by Polajnar *et al.*, 2015; Durel *et al.*, 2015), insect pesticide resistance, and a better appreciation of the associated external costs of pesticides (Leach & Mumford, 2008), drives the need for alternative and more sustainable (low carbon) control strategies (Arthur, 1996). Adding global warming into the mix introduces another element of uncertainty to the future of plant protection practices (Godfray *et al.*, 2010; Yi *et al.*, 2014). Global trade has exacerbated problems further by aiding pest species migration and dispersion (Waage & Mumford, 2008).

Accordingly, integrated area-wide approaches to pest management are desired and may offer a more sustainable future for agriculture (Pimbert, 1991; Van Lenteren, 2000; Pimentel, 2007). Perhaps this century's food crisis will be solved by an agricultural paradigm shift, from the Green Revolution, towards less intensive (fewer inputs) yet equally productive sustainable agriculture (Pretty, 2008), confounding Malthus once again.

1.1.2 POSSIBLE SOLUTIONS – AN INTEGRATED APPROACH TO AGRICULTURE

One possible avenue of research for sustainability relies upon the biotechnological innovations and associated insect genomic manipulations of the last 15 years (Fraser, 2012; McGraw & O'Neil, 2013). The use of genetically modified organisms (GMOs) in agriculture is not new (Ronald, 2011). Historically, the majority of applications of GMO have been in the form of transgenic plants (Hokanson *et al.*, 2013), such as *Bacillus thuringiensis* (*Bt*) cotton or glyphosate (N-(phosphonomethyl)glycine)-resistant crops, with genetically modified (GM) crops currently being grown in an area greater than 134 million hectares (Ronald, 2011; Qaim & Kouser, 2013). Some proponents hail these developments as the beginning of the Gene Revolution (Wu & Butz, 2004), albeit with the same hurdles and issues to overcome as a decade ago (Atkinson, 2002).

The genetic tools and methods made available from these advances facilitate improvement in pest suppression options (Alphey *et al.*, 2006). This thesis will focus on specific facets of genetic control of agricultural insect pests as a solution; which may be defined as “Dissemination, by mating or inheritance, of factors that reduce pest damage” (Alphey, 2014). This modern approach to pest management was preceded by the traditional sterile insect technique and is a useful tool in the pest manager’s arsenal.

1.1.3 THE STERILE INSECT TECHNIQUE

The sterile insect technique (SIT) was conceived by E. F. Knipling (1955), A. S. Serebrovskii and F. L. Vanderplank (Klassen & Curtis, 2005). The method relies on inundating a wild pest population of a given species with mass-reared sterilised males of the same species. For SIT to be effective, sterile males need to locate, attract, court, copulate, effectively inseminate, and inhibit wild female re-mating. Sterile male fulfilment of this checklist reduces the number of viable matings between wild males and females in the target area.

Repeated releases of sterile males can suppress population growth and can result in eradication in a species-specific manner (Krafsur, 1998). Conventional SIT employs ionic radiation (or chemosterilants) to cause chromosomal aberrations or double-stranded breaks in the germ line deoxyribonucleic acid (DNA) of the reared males (Helinski *et al.*, 2009). This results in the early death of sired zygotes, due to lethal imbalances of genetic material, when the mitotic cleavage division occurs. In effect, appropriately dosed insects can be considered “sterile”.

SIT has been successful in controlling diverse agricultural pest species (reviewed in Morrison *et al.*, 2010) since the 1950s with cost-benefit ratios as high as 1:10 (New World Screwworm, *Cochliomyia hominivorax*) with some eradication programmes (Klassen & Curtis, 2005).

These successes have been dependent on concomitant improvement in diverse disciplines associated with area-wide SIT (AW-SIT). These developments include methods of improving the quality of mass-reared insects (Simmons *et al.*, 2010), tools for thorough cost-benefit analysis

(Mumford, 2005), administrative improvements including programme management (Dyck *et al.*, 2005), public-relations and political support consideration (Dyck *et al.*, 2005), advantages of modern geographic information systems (GIS) (Cox & Vreysen, 2005), and an increase in the number of successful case-studies (Hendrichs *et al.*, 2005). SIT is not a stand-alone technology, and thus it has to be integrated with other control tactics and knowledge-intensive methods.

Given the advantages of SIT and the success stories, why has it been implemented against so few major pest species? This phenomenon can be explained away, rather glibly, by constraints and misconceptions, or more often, by a combination of constraints that are biological, financial, social or political in nature (Whitten & Mahon, 2005).

A common misconception is that there is only one type of SIT. There are many flavours of SIT that have been developed to solve specific pest problems. Going forward it is important to take a historical perspective to appreciate lessons learned in over six decades of genetic control of insect populations (Black *et al.*, 2011). Comparing some of the older methodologies for introducing genetic change or rendering males effectively sterile with modern technological advances foreshadow a century of future SIT-like successes (Burt, 2014).

1.1.4 SIT BY GENETIC MODIFICATION

The Gene Revolution has provided the knowledge and tools to use genetic techniques to provide effective sterility (not true sterility, as with irradiation) in a range of pest species in which SIT is not used, or hitherto was deemed non-viable, or to improve existing SIT programmes. Such genetic techniques are developed by inserting foreign genetic constructs into the host DNA (Handler, 2000). For heritable genetic changes, the desired genes are inserted into the germline, as only these types of cell can form gametes and contribute to the genotype of the next generation of organisms (Clark & Pazdernik, 2012). These genetic changes are introduced by co-opting naturally occurring transposable elements or “jumping genes” discovered by McClintock (1950).

With the first transposable element (transposon)-based insertions of a transgene into an insect genome, in *Drosophila melanogaster* (Rubin & Spradling, 1982), research attention heightened in the area of insect transgenesis. However, progress slowed when it was found that the *P*-element used in the *Drosophila* transformation was not transferable to other insect species (Morris *et al.*, 1989). Transformation of other species has therefore relied on alternative transposable elements, such as *piggyBac* and *mariner* (O'Brochta & Atkinson, 1996).

Successful transformation has been achieved in a range of economically important pest species using transposable element-based vectors (O'Brochta & Atkinson, 1996; Thresher *et al.*, 2009; Kim & Pyykko, 2011). One of the most commonly used of these is the *piggyBac* element, originally discovered in cell lines of the cabbage looper moth, *Trichoplusia ni* (Fraser *et al.*, 1983; Sarkar *et al.*, 2003; Robinson *et al.*, 2004; Zimowska & Handler, 2006). It has been used successfully for germline transformation in multiple insect orders: Diptera (Labbe *et al.*, 2010), Lepidoptera (Martins *et al.*, 2012), Coleoptera (Lorenzen *et al.*, 2007; Wang *et al.*, 2008) and Hymenoptera (Sumitani *et al.*, 2003).

A recent review by Kim & Pyykko (2011) summarises the molecular structure and mobility of *piggyBac*. The 2472-bp-long element is structured with two sets of inverted repeats at both ends and a central transposase encoding open reading frame (Fraser, 2012). The insertion site of *piggyBac* is quasi-random, with a cut-and-paste insertion at the short genome motif site of TTAA (O'Brochta *et al.*, 2003; Wu & Burgess, 2004; Zhuang *et al.*, 2010). A more precise assessment of where the *piggyBac* inserts into the host genome with respect to chromatin structure revealed hotspots and cold spots in the *Drosophila melanogaster* genome (Bellen *et al.*, 2011) (see also Li *et al.*, 2013). The site selectivity of *piggyBac* suggests it has evolved to reduce deleterious and increase adaptive changes in host gene expression, mitigating insertional effects compared to the use of other transposons for transgenesis (Figure 1). However, there is evidence that unintended traits do not always result following an insertion (Scolari *et al.*, 2011; Schnell *et al.*, 2014).

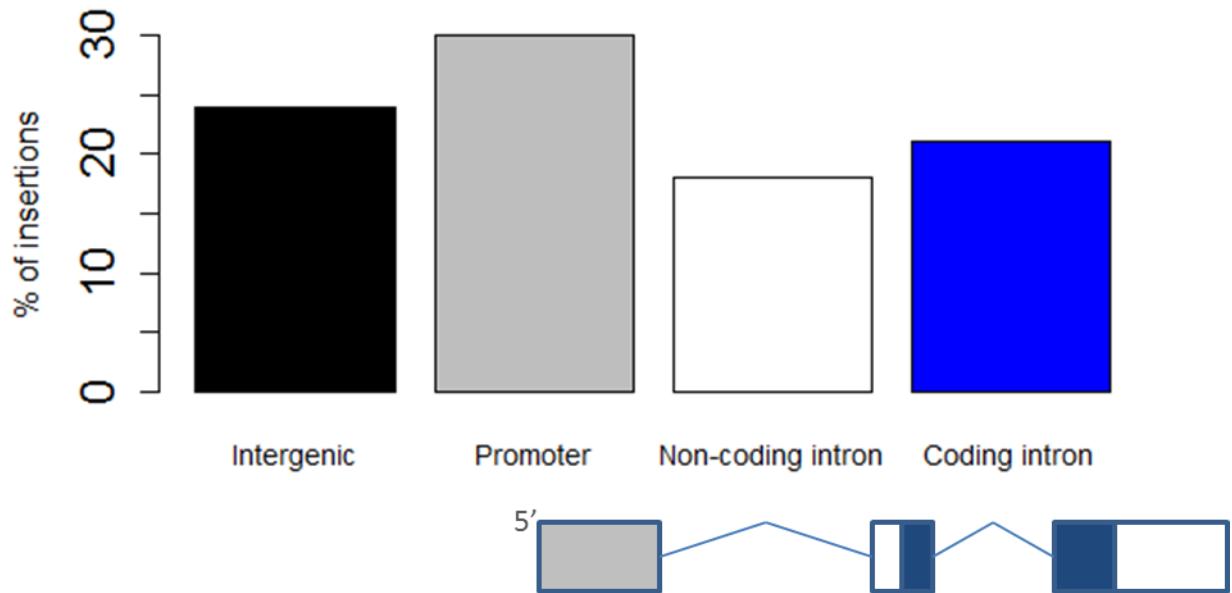


Figure 1. *piggyBac* insertion with respect to transcript structure – the percentage of insertions located in the indicated regions of annotated transcripts, after Bellen *et al.*, 2011. Based on 12,244 insertions; as part of the *Drosophila melanogaster* gene disruption project. Intergenic insertions are those > 500 bp upstream of the transcription start site. Note that in this experiment the transgene moved from one location in the DNA to another not from a plasmid to chromosomal DNA, as is the case with insect germline transformation. Non-coding introns are introns found in non-coding RNA genes. Coding introns are introns found in genes that code for protein.

The advantages of *piggyBac* for insect transgenesis include:

- *piggyBac* can carry up to 14 kb of foreign genes
- acceptable genomic insertion efficiency
- precise integration and excision
- pseudo-random integration at TTAA sites
- post-integration stabilisation possible
- historical use and reliability

Inevitably, a “jumping gene” could cause problems concerning regulation of the technology (Reeves *et al.*, 2012). To prevent post-transformation transgene mobility (Sethuraman *et al.*, 2007) the open reading frame (ORF) of the transposase is typically removed from the transformation construct, to render it non-autonomous (Condon *et al.*, 2007). Remobilisation of

the integrated construct is infrequent (O'Brochta, 2003), and transposase is still necessary for the insertion event.

A separate helper (DNA plasmid or mRNA) encoding the transposase can be provided at the time of transformation, for example during microinjection of embryos, to permit transposition (Shinmyo *et al.*, 2004). The transposase allows the transgene to be cut and pasted into the host genome (Mitra *et al.*, 2008). Plasmid DNA is the most commonly used method of delivery of a transposase source as it is stable, mRNA helper is also viable (Kapetanaki *et al.*, 2002). If DNA is used the choice of constitutive promoter is relevant to transposase transcription and transformation efficiency (Li *et al.*, 2001). The amino acid sequence of the transposase used has also been shown to affect transformation efficiency (Wright *et al.*, 2013). Other factors may affect transformation efficiency, which is relevant to a researcher when deciding how many injections to conduct.

The relationship between transposase concentration and transposition rate in *piggyBac* is equivocal. A recent review has shown inhibition (of transposition) by transposase overconcentration (ITOC) in some cell lines with variation in mode of regulation between somatic and germ line cell types. This uncertainty is compounded by incomplete methods descriptions in the literature where the concentration of helper plasmid used is not specified (Martins *et al.*, 2012). Thus the optimum concentration of helper plasmid with which to co-inject is often unknown.

The vector and helper plasmid are injected into the posterior end of the egg, with the aim that the transposon system will cut and paste the construct into the genome of a germline pre-cursor cell. Circular versions of the DNA construct vector have higher efficiencies compared to linear alternatives (Nakanishi *et al.*, 2011). If integration into the germ line occurs a proportion of progeny will be transgenic. For integration to be detected a coding sequence for a reporter protein-based signal is required.

Transposon sequences are used in combination with marker systems for detectable insect transgenesis. A range of marker systems are available (Jones *et al.*, 2012). Visible fluorescent proteins markers are the most commonly used methods in pest insects, and are readily transferable between species (Berghammer *et al.*, 1999). Enhanced green fluorescent protein (Verkhusha & Lukyanov, 2004; Shagin *et al.*, 2004) (ZsGreen), red fluorescent protein drFP5833 (DsRed) (Bevis & Glick, 2002) and enhanced cyan fluorescent protein (ECFP, e.g. AmCyan) are all examples of proteins that have been modified from naturally occurring wild-type proteins to improve absorption and quantum yield (Patterson *et al.*, 2001).

For SIT-type approaches, these fluorescent protein markers offer an additional advantage (Schellhorn *et al.*, 2004), providing a method for monitoring relative numbers of transgenic and wild insects in the field. Such an approach has been demonstrated in the pink bollworm, *Pectinophora gossypiella* (Simmons *et al.*, 2011; Morrison *et al.*, 2012).

1.1.5 TARGET SPECIES

This thesis is targeted against two insect orders that are underrepresented in the AW-SIT control world. SIT has been underutilised against the Lepidoptera and the Coleoptera with only three and one operational control programmes attempted, respectively.

Specifically, in the Lepidoptera, pink bollworm (*Pectinophora gossypiella* Saunders) in the USA, a major cotton pest; codling moth (*Cydia pomonella* L.) in Canada, a pest of apples and pears; and painted apple moth (*Orgyia anartoides*, Walker) in New Zealand, a polyphagous forestry and horticulture pest (Bloem *et al.*, 2005).

SIT against the Coleoptera has been successfully trialled against the cockchafer (*Melolontha vulgaris*, F.), boll weevil (*Anthonomus grandis grandis*, Boheman) and, the only recently- active programme, attempting control of the sweet potato weevil (*Cylas formicarius*, F.) in Japan (Klassen and Curtis, 2005).

These short lists contrast to the number of operational programmes focused on control of Diptera; this thesis seeks to address this skew by developing genetic control strategies in two important agricultural model-pest species.

1.1.5.1 DIAMONDBACK MOTH

The diamondback moth (DBM), *Plutella xylostella* (L.) is a major pest of cruciferous crops, with a worldwide distribution (Talekar & Shelton, 1993). Diamondback moth is economically damaging with estimated direct and indirect costs of US\$4-5 billion per annum (Zalucki *et al.*, 2012). The larval stages feed on foliage and can cause significant damage with up to 90% yield losses (He *et al.*, 2012). The cryptic feeding habits and rapid generation times render it difficult to control. Talekar and Shelton (1993) describe a range of biological and ecological characteristics that contribute to its pest status, also blaming uncontrolled pesticide application as a reason for the species developing resistance to pesticides so rapidly (Scott, 2008; Bravo & Soberon, 2008). The pest is capable of both short- and long-range dispersal (Mo *et al.*, 2003).

The genome has been sequenced independently by three groups based in Japan (<http://dbm.dna.affrc.go.jp/px/>), China (<http://www.iae.fafu.edu.cn/DBM/index.php>) and Liverpool (http://www.cgr.liv.ac.uk/gview/plutella_test) respectively. I utilised the Liverpool data for this project, as the laboratory strain in our laboratories was the subject of this sequencing work (unpublished, Dr Alistair C. Darby, Centre for Genomic Research, University of Liverpool, UK). Hereafter these genome data sources will be referred to as Japan, China or Liverpool sequence data.

Cataloguing genes that are over-expressed in one sex relative to the other in sexually dimorphic tissue can help identify candidate genes and pathways responsible for sexually dimorphic development (Oppenheim, 2015). To facilitate the design of a paternal effect system, RNA-seq data was also required comparing unique transcript abundance in testes to non-testes tissues. The availability of a decent genome scaffold facilitated transcriptome assembly and annotation.

1.1.5.2 RED FLOUR BEETLE

The stored product pest, the red flour beetle (*Tribolium castaneum*, Herbst), shares diamondback moth's propensity for developing resistance to insecticides, and has proven similarly difficult to control by conventional insecticide-based means (Sokoloff, 1972).

Tribolium are thought to have occupied an ecological niche feeding on underground animal food stores in arid sub-tropical regions; 'pre-adapting' the species for a commensal relationship with humans going back to at least 2000 BC (Fedina & Lewis, 2008). This evolutionary background has provided *Tribolium* as an excellent research organism as it is easy to rear and has a short life-cycle (Sokoloff, 1972). The recent sequencing of the *Tribolium* genome (<http://beetlebase.org/>) greatly facilitates genetic research on this pest and other beetles (Richards *et al.*, 2008).

1.1.6 LIMITATIONS OF THE TRADITIONAL SIT APPROACH

Conventional SIT relies on irradiation to sterilise the target pest species. However, the ionising radiation required to induce sexual 'sterility' in insects often impairs performance to such an extent that the males are not effective in an SIT programme (Bloem *et al.*, 2004; Helinski *et al.*, 2009). Genetic approaches seek to overcome these problems, but the associated insertional mutagenesis of the transposition event and the physiological disruption of the construct itself can affect performance also (Catteruccia *et al.*, 2003; Marelli *et al.*, 2006; Marelli *et al.*, 2006; Bargielowski *et al.*, 2011; Scolari *et al.*, 2011). The mass rearing process alone has been shown to adversely affect male quality but can be compensated for by releasing at high flooding ratios (Boake *et al.*, 1996). Poor competitiveness of released insects can be a major constraint on the success of a SIT programme (Calkins & Parker, 2005).

There is not always a viable sexing mechanism appropriate for mass-rearing (Papathanos *et al.*, 2009), resulting in both male and female insects being reared within the programme facilities.

This increases production costs and may decrease SIT efficiency (Alphey *et al.*, 2008) as sterile females have been shown to 'distract' co-released sterile males from seeking wild females in some species (Rendon *et al.*, 2006).

SIT does not entail any intellectual property rights (IPR); accordingly there is little incentive for private enterprise to participate, with the profits coming exclusively from the sale of services (Whitten & Mahon, 2005). Genetic control, on the other hand, provides additional value to investors, where even if the venture were to fail, the control of the associated IPR could mitigate the financial risk.

Ultimately, any SIT programme will need to be judged in terms of its costs and benefits (Mumford, 2005; Morris, 2011), regardless of the methods used to provide male sterility and the concurrent population suppression. Perhaps this is easier said than done, with the efficacy of SIT often confounded by other treatments that are used concurrently (Krafsur, 1998).

1.1.7 RELEASE OF INSECTS CARRYING A DOMINANT LETHAL

A variant of the SIT, known as RIDL (Release of Insects carrying a Dominant Lethal), was first demonstrated in *Drosophila melanogaster* (Thomas *et al.*, 2000) and later in a number of pest species (Gong *et al.*, 2005; Phuc *et al.* 2007; Fu *et al.* 2007), provides a promising means of overcoming the limitations of SIT, while retaining the key advantages. A comparison between the different genetic control strategies is detailed in Black *et al.* (2011).

Like conventional SIT, RIDL does not produce sterile insects *per se*; rather, the modified organism carries a conditional, dominant gene (Black *et al.*, 2011). The permissive conditions are provided by a chemical additive that is not encountered by the insects in the wild but can be supplied to the organisms in the rearing facilities in their diet (for example, tetracycline or suitable analogues) (Gossen & Bujard, 1992).

Desirable characteristics of an idealised binary-system for this role include; silent expression in the presence of repressor (low expression leakiness), high expression in the absence of repressor, high specificity and sensitivity to absence of repressor, quick response to absence of repressor, regulation by an orally bioavailable repressor, minimal or no immune impact to the host and finally *in vivo* applicability (Sotiropoulos & Kaznessis, 2007). The most widely used inducible transcription systems that largely meet these criteria are the tetracycline regulatory expression systems based on the tetracycline resistance operon of *Escherichia coli*.

This is one of many binary expression systems implemented in transgenic insects (Viktorinova and Wimmer, 2007). The Tet-Off system has the advantage of strict temporal control, in contrast to some other binary systems such as the Gal4/upstream activation system (UAS). The Tet-Off system uses tetracycline-controlled transactivator (tTA) and its tetracycline operator sequence binding site (*tetO*). The tTA element is a fusion protein combining the tetracycline repressor (tetR) from *Escherichia Coli* with the activating domain of the herpes simplex virus protein 16 (VP16) (Hara *et al.*, 2009) (the fusion is called tTAV). tTAV is a variant tTA sequence optimised for expression in *D. melanogaster*, while giving only low basal expression when exposed to tetracycline (Gong *et al.*, 2005).

The tTAV is regulated by a promoter in the first module of the construct. The second module contains the target gene under the control of seven copies of the bi-directional tetracycline operator (*tetO*) and a minimal promoter. In the absence of tetracycline, tTAV binds to the *tetO* as a dimer, and transcription of the effector gene is enhanced (Figure 2). In the presence of tetracycline a conformational change in the tetracycline repressor prevents the tTAV from binding the DNA, thus repressing the enhancement provided by the *tetO* on the minimal promoter (Gallia & Khalili, 1998).

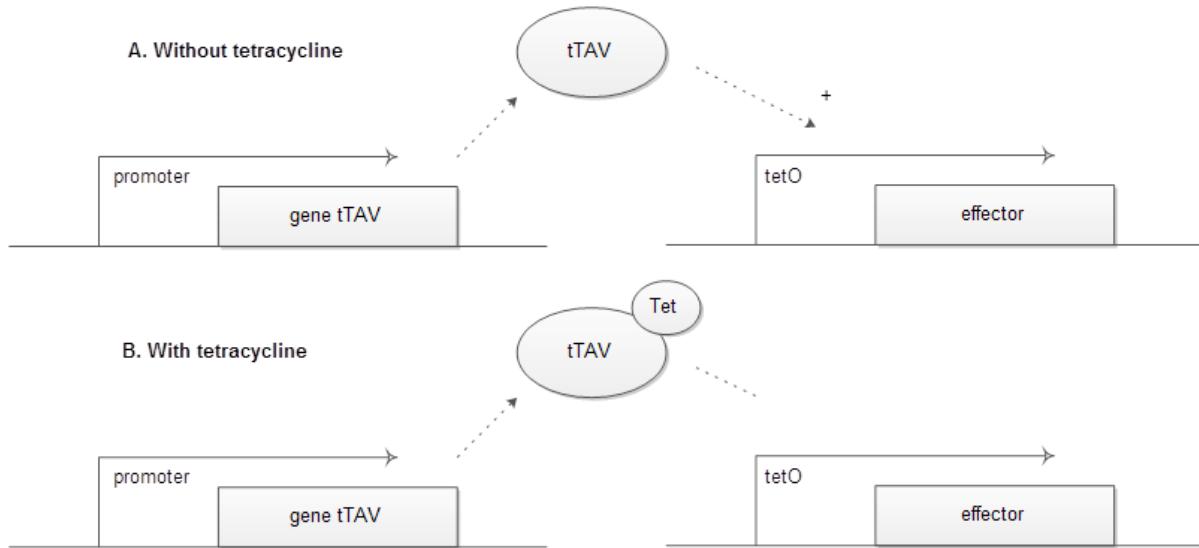


Figure 2. A suitable promoter (i.e. sex, life-stage specific depending on objectives) controls the production of the transcriptional activator (tTAV) protein. This protein can initiate transcription of the tetO regulated effector (typically a minimal promoter such as *Dmhsp70* is included between the tetO and effector). (A) Upon translation the tTAV protein binds to the tTAV response element (*tetO*), enhancing further expression of the lethal effector, and leading to death. (B) Upon translation the tTAV binds preferentially to tetracycline ('Tet') instead of *tetO*, repressing minimal-promoter enhancement and subsequent lethality.

This system is amenable to modification and can be changed to meet specific requirements including auto-regulatory control. A one-component positive feedback loop has been effective as part of the RIDL system. In the absence of tetracycline, basal expression of tTA leads to the synthesis of more tTA (the *tetO* and a minimal promoter drive tTA expression), which accumulates to a high level resulting in subsequent death of the cell (Gong *et al.*, 2005). The ideal is for the effector to be responsible for 100% lethality (complete penetrance) in the absence of tetracycline and 0% lethality in its presence.

If the promoter driving the tTAV expression is active in both sexes then it results in bi-sex lethality. This form of RIDL is analogous to conventional SIT, and if male-only releases are preferred females have to be removed mechanically. Recent trials with the OX513A tetracycline-repressible late-acting dominant lethal *Aedes aegypti* line, in which mechanical sieving was used to separate the sexes as pupae, have been effective in the field (Harris *et al.*, 2012; Alphey *et al.*, 2014).

1.1.8 SEX SORTING

Separating sexes manually can be time consuming, expensive and is not possible in some species. A repressible genetic sexing mechanism (GSM) that causes female-specific lethality off tetracycline (fsRIDL) offers a genetic solution. This can and has been achieved using a variety of strategies. From exploiting sex-alternate splicing (Fu *et al.*, 2007; Ant *et al.*, 2012), producing functional tTAV only in the females, to use of a female-specific promoter to drive expression of tTAV (Fu *et al.*, 2010). Bi-sex and female-specific lethality can be compared using truth tables (Table 1).

Table 1. The transgenic construct provides a genetic circuit that responds to inputs of sex and presence of tetracycline in the larval diet and gives an output of death when activated. This is as a truth table where an output value of 1 represents a viable condition and 0 a lethal condition. Bi-sex RIDL insects are non-viable off tetracycline. Female-specific RIDL females (males are viable) are non-viable off tetracycline (after Jin *et al.*, 2013).

| Bi-sex RIDL | | | fsRIDL | | |
|--------------------|------------------------|-----------|---------------|------------------------|-----------|
| Tet | Sex | | Tet | Sex | |
| | σ^{\rightarrow} | φ | | σ^{\rightarrow} | φ |
| + | 1 | 1 | + | 1 | 1 |
| | 0 | 0 | | 1 | 0 |

Interestingly, upon mating a wild-type female, a transgene-homozygous fsRIDL male will sire female and male offspring. The F₁ female progeny die with the male heterozygotes surviving to pass on the fsRIDL transgene to the following F₂ generation (Figure 3). Thereafter, the gene leaves the population analogous to radioactive decay due to associated fitness costs (Harvey-Samuel *et al.*, 2014).

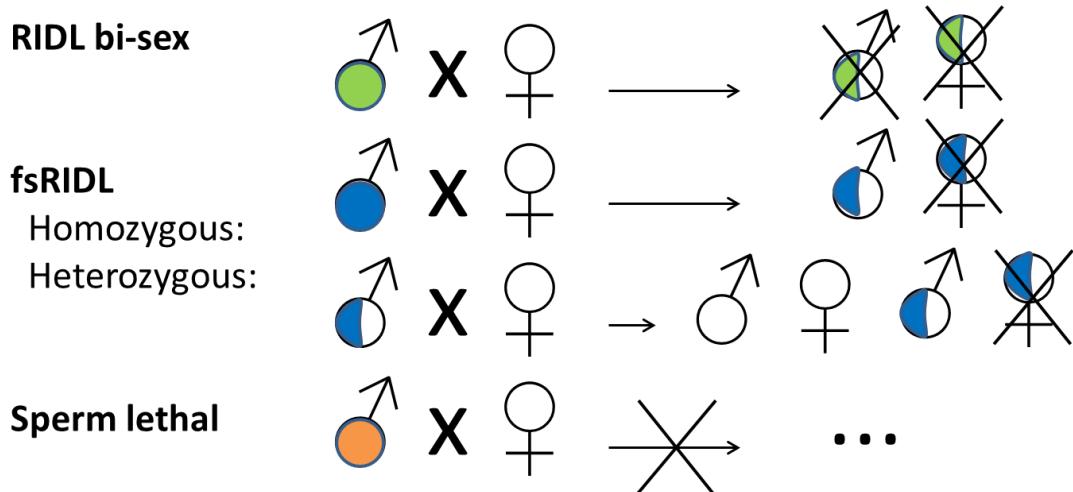


Figure 3. Selected genetic strategies. RIDL, fsRIDL and paternal effect involve the release of transgenic males homozygous (filled symbols) for a repressible dominant gene to mate wild females (open symbols). For bi-sex RIDL (green), as for SIT, the heterozygous (half-filled symbols) offspring are killed by the inherited lethal gene without access to the antidote Tetracycline. In fsRIDL, the lethal gene (blue fill) is female-specific, so daughters die but heterozygous sons live; if these mate, half of their offspring inherit the female-lethal transgene. The sperm lethal system is better suited to agricultural pests as the homozygous male mating with wild type female does not produce any viable offspring due to double-stranded breaks in the male sperm DNA (after Alphey, 2014).

The paternal effect system is introduced above. Interestingly it is more similar to conventional radiation-treated sterile males in that no viable offspring are produced, hence the crossed arrow in Figure 3. Depending on the pest species in question, one method may be preferred over the others.

1.1.9 CONVENTIONAL RIDL VERSUS PATERNAL EFFECT

1.1.9.1 BENEFITS OF A PATERNAL EFFECT SYSTEM

The RIDL system is dependent on a lethal effector gene, or lethality caused by tTAV over-expression or “transcriptional squelching” (Lin *et al.*, 2007), which could theoretically be disrupted by mutation. Computer modelling by Alphey *et al.* (2011) has identified the plausibility and risk of resistance to this genetic control system emerging by the evolution of: resistance to the lethal effector; an antidote to the lethal effector; or a tTAV protein inhibitor. Hence a system that introduces double-stranded breaks into the gamete DNA, analogous to the

radiosterilisation method, would be preferred as the evolution of resistance to such breaks appears highly improbable (although history has taught us to be careful making such claims).

The development of such a system would be reliant on the expression of endonucleases during a specific window of sperm development to create the double-stranded breaks, resulting in a sterile GMO, a boon for overcoming regulatory restrictions and a means to extend patent life, driving external investment and the actualisation of genetic control in the field. Accordingly such a system includes several traits, some of which also apply to SIT and RIDL. In this way the paternal effect system can be considered third generation technology, with bisexual RIDL as first- and fsRIDL as second-generation (Table 2).

Table 2. Relative benefits of three generations of insect genetic control systems at Oxitec for population suppression or eradication. The + describes the truth of the consideration or the relative ease of that consideration relative to the other strategies which will have a – if they fail in that consideration or are weaker in some aspect. The table attempts to tease out the differences between the generations of the technology, with first second and third from left to right. The binary nature and limited detail may ignore some of the finer nuances of the more important points which are elucidated in the main body of the text.

| Genetic strategy | | | |
|---|----------------------|-----------------------------|------------------------------|
| Considerations | 1. Bisex RIDL | 2. fsRIDL | 3. fs-paternal effect |
| Published example | OX513 | OX4319 | Patent pending |
| Species-specific | + | + | + |
| Resistance management | + | ++ | + |
| As part of an IPM approach | + | + | + |
| Genetic removal of factory females | - | + | + |
| Conditional | + | + | + |
| Orthogonal systems | possible | possible | possible |
| Resistance risk | unlikely | unlikely | Very unlikely |
| Transgenic insects ingested by humans | Possible, adults | Possible, larvae and adults | Possible, as eggs |
| Oviposition damage risk from mated wild-females | + | + | Possibly not |
| Chance of phenotypic rescue due to environmental antidote | + | + | - |
| Technology transferability between species | + | - | - |

Agricultural pests generally cause most economic damage as larvae, directly due to feeding and indirectly due to acting as a vector for pathogens (Savary *et al.*, 2006). This results in yield loss for the grower as well as affecting the quality of the surviving crop. For some crops the relationship between pest damage and loss of value of a crop can be non-linear, such as in the crucifers (Subramanian *et al.*, 2010). Unlike RIDL and fsRIDL, the paternal effect system can be used to kill fertilised embryos early on in development. This prevents larval feeding and may

have the added benefit of repressing oviposition by the mated wild-female. This is the core benefit of the system and highlights its utility when applied to agricultural pests.

The paternal effect system is designed to introduce irreparable damage to the DNA of the sperm but at a point in development where the normal sperm phenotype is unaffected (similar to conventional-irradiation dominant-lethal mutation inducing based SIT). When a break is induced in a chromosome in mature sperm, it remains in this condition until after the sperm has entered an egg. Following fusion, nuclear divisions begin, and a break in a chromosome has drastic effects on the embryo viability leading to the breakage-replication-fusion-bridge cycle and eventual death (Robinson, 2005). However, some Lepidoptera have proven resistant to conventional irradiation putatively due to the holokinetic nature of the chromosomes (irradiation has historically been successful in insect species with the more typical monocentric chromosomes; see Figure 4 for the difference).

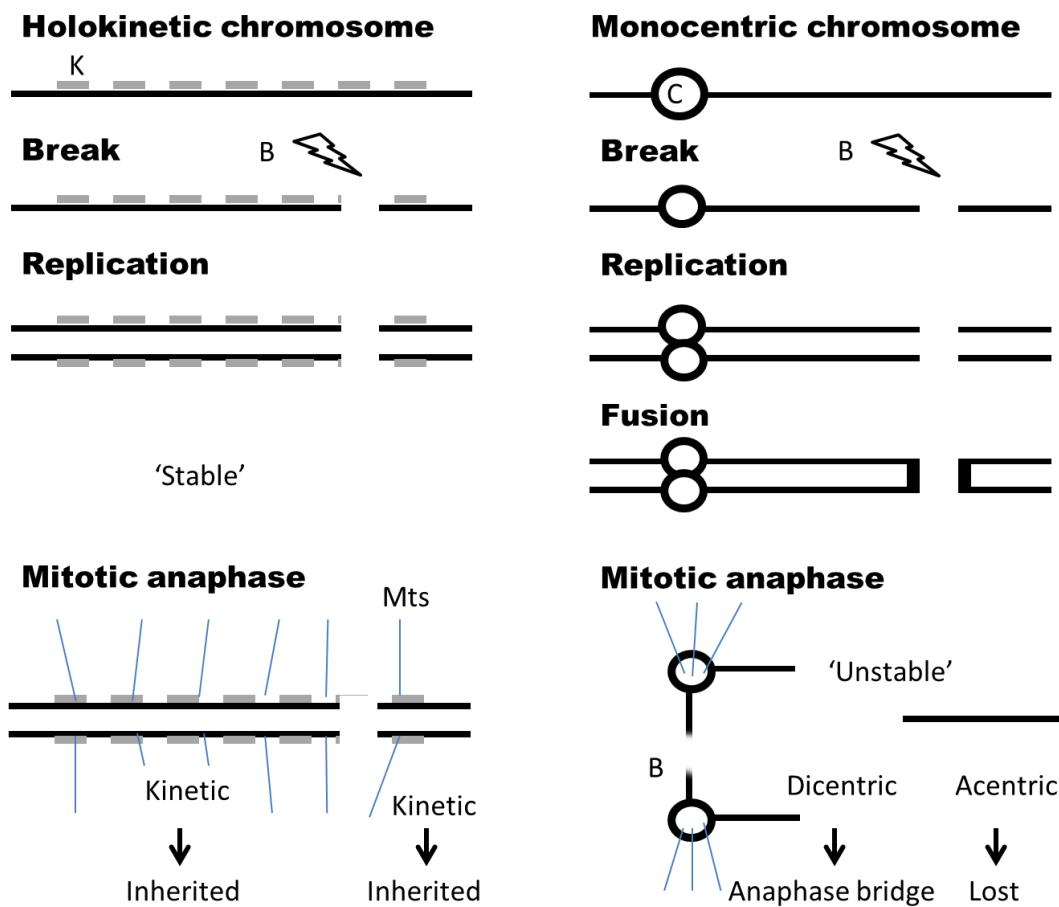


Figure 4. The mitotic fate of chromosomes with induced dominant-lethal mutations in a holokinetic (typically lepidopteran; left) chromosome and a monocentric (typically dipteran; right) chromosome. In the holokinetic diagram the kinetochore plate (K) extends across more than 50% of the chromosome. If a breakage (B) occurs within K, the resulting fragments each contain points for spindle microtubules (MT) to attach (are kinetic) and will be inherited correctly. For the monocentric chromosome a strand breakage results in an acentric fragment (without a centromere (C)) and a dicentric bridge. The ends of these fragments fuse (F). During anaphase, the acentric fragment will be lost while the dicentric fragment will form a bridge, causing further chromosome breaks. The breakage-replication-fusion-bridge cycle causes death. Modified from (Carpenter, Bloem et al. 2005; after Harvey-Samuel thesis, 2014).

Paternal effect may offer a solution to this problem of radio-resistance in the Lepidoptera.

Nuclease expression in the sperm, as part of a paternal effect system, could fragment the DNA at sufficiently frequent intervals to cause gene rupture or uneven distribution of genetic material into daughter cells during mitosis after fertilisation, despite molecular repair mechanisms.

This design is preferred to producing spermless males, as some insect species altered in this way upon mating with females do not instil the desired behavioural copulatory refractoriness (Ravi Ram & Wolfner, 2007). Counter to this it has been shown to be effective in at least one

species (Thailayil *et al.*, 2011). Ultimately, the probability of wild female mating failure (Perez-Staples *et al.*, 2013) determines the success of the control strategy.

1.1.9.2 INTEGRATED GENETIC CONTROL AS PART OF AN INTEGRATED PEST MANAGEMENT APPROACH

These strategies are not mutually exclusive. The advantage of paternal effect combined with fsRIDL over fsRIDL alone comes to the fore when discussing agricultural pests where both male and female larvae cause unacceptable feeding damage. The development of the paternal effect system would provide the benefit of no larval feeding damage while the fsRIDL transgene component could provide all male releases by repressible female mortality. This is in contrast to mosquitoes, whereby late-acting lethality is beneficial due to density-dependent competition effects (Phuc *et al.*, 2007). Accordingly, early-acting lethality is preferred as even female oviposition can cause secondary infections in crop plants.

Ultimately this thesis is about striving towards the next generation of genetic control of agricultural pests, the culmination of over a decade's worth of work (Figure 5).

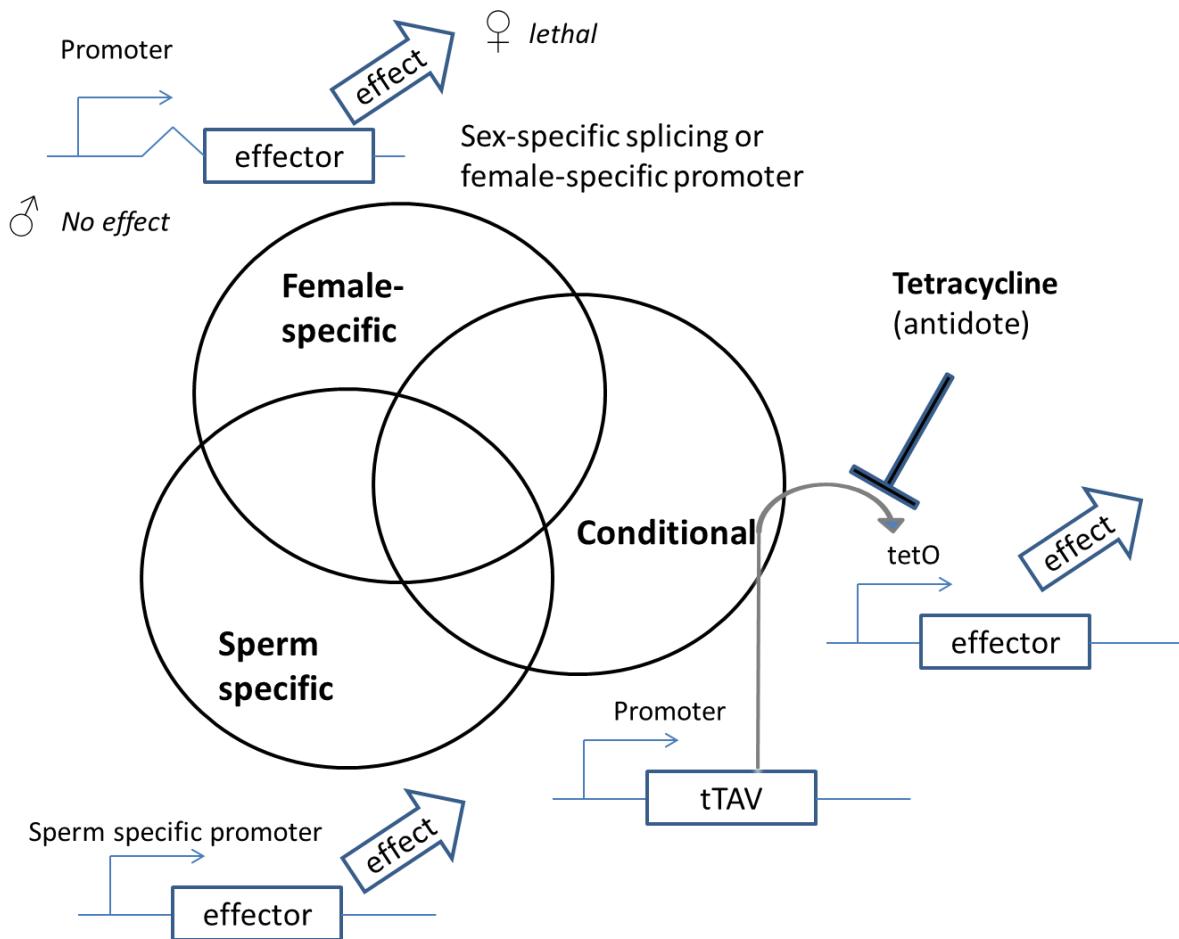


Figure 5. The paternal effect system as applied synthetic biology: combinatorial control of gene expression with defined modular control elements. In the constructs described here, the sperm-specific promoter, sex-specific intron (wavy line) and tetracycline-repressible trans-activation mechanism (tTAV or alternatives bind and activate the tetO enhancer in the absence of tetracycline) act as independent control elements; logic gates which combine specific inputs (respectively Tissue and / or time, Sex, Conditionality) to give predetermined logical outputs (After Labbe *et al.*, 2012). The optimal design depends on the characteristics of the pest species in question. Mosquito control requires a different solution to leaf-miner control, for example.

Currently the Tet-Off system is preferred as a robust means to ensure conditionality. However, the use of an orthogonal system alongside would provide additional fine tuning to ensure optimal timing and efficiency of the operation of each module (Gitzinger *et al.*, 2012; Folcher *et al.*, 2013). For example, it may be preferred to provide the antidote for the female-specific system earlier than for the sperm-specific module.

1.1.10 MATHEMATICAL MODELS FOR THE APPLICATION OF STERILE INSECTS

SIT has been a fruitful playground for the development of mathematical population models with both analytical and numerical approaches to predicting the behaviour of a pest population applied historically. Predictive models, based on empirically derived parameter estimates, can be used to determine the dynamics of a population or various features thereof. For example a simple growth model describes how a population increases exponentially or a logistic model which curbs the population growth to the carrying capacity of the environment, affecting the change in the population growth rate due to density-dependent growth (Kimanani & Odhiambo, 1993).

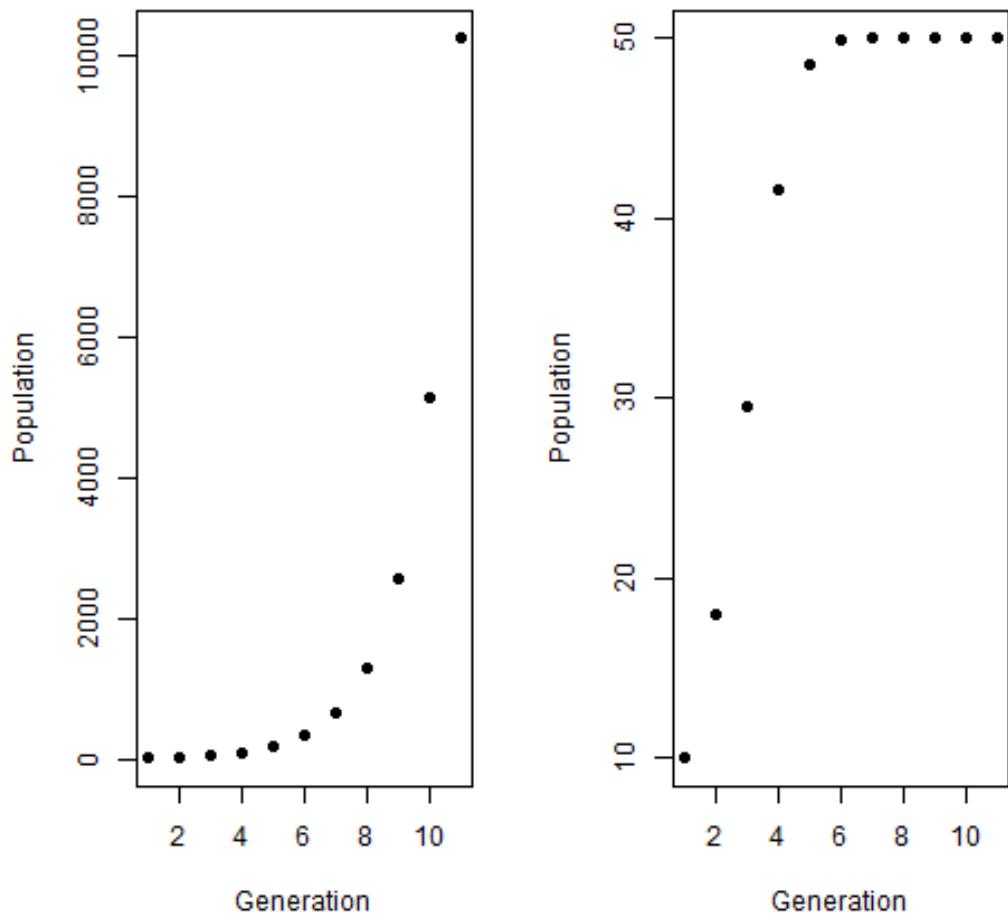


Figure 6. Simple population models over ten generations showing the exponential growth assumed when there is no density dependence (left) and there is density dependence (right). Notice the different scales on the y-axis. The reproductive rate, $a = 2$, the carrying capacity $k = 100$, the initial population $y_0 = 10$. The functions were written in R.

Knipping built on these earlier simple population models by providing a simple numerical model. The central novel feature was the ratio of fertile males to all males in the population ($M/(S+M)$) where M is the number of fertile males and S is the number of sterile males (Dyck *et al.* 2005).

For example if we used the logistic population model (the right) but introduced sterile males at an over-flooding ratio of 10:1 ($S:M$) we would see a decline in the pest population. In this situation at generation eleven, the wild population would be comprised of 25 M and 25 F. Our release of 250 S would reduce the chances of a F and M mating to 9%. Ergo only 9% of F

produce offspring (2 per female). If we introduce the additional assumption that F (and M and S) dies after the mating period then the population will decline to 18. At this rate the population reaches extinction at generation 13.

This type of modelling elucidates a very important attribute of SIT. In contrast to traditional chemical spraying, SIT provides more effective population suppression at lower pest densities. Eradication of a pest species from an area can and has been achieved as part of an integrate eradication programme (Suckling *et al.*, 2014).

These simple models have been extended to build more complicated tools used today comprised of many parameters (Seirin Lee *et al.*, 2013). These models are useful for informing decision makers on the optimal life stage for lethality, over-flooding ratios and frequency of release. Often the tools can be co-opted and applied to subtly different scenarios. For example this thesis is focused on the sperm lethal system whereby S compete with M for the chance to mate with females (F). The putative offspring, when S are successful, will be non-viable.

1.1.11 CONSTRUCT DESIGN

Several factors are important when designing constructs for transformation of pest species. The phenotype of transformed insects is determined by several factors including: the regulatory elements that drive transgene expression; the effector(s); and the transgene insertion site (Franz *et al.*, 2011) (Figure 7). The insertion site cannot be controlled with *piggyBac* but we can control transgene expression using a variety of exogenous and endogenous genetic components, depending on the specific functional requirements. The primary marker and promoter combination is also crucial for detecting successful transformation.

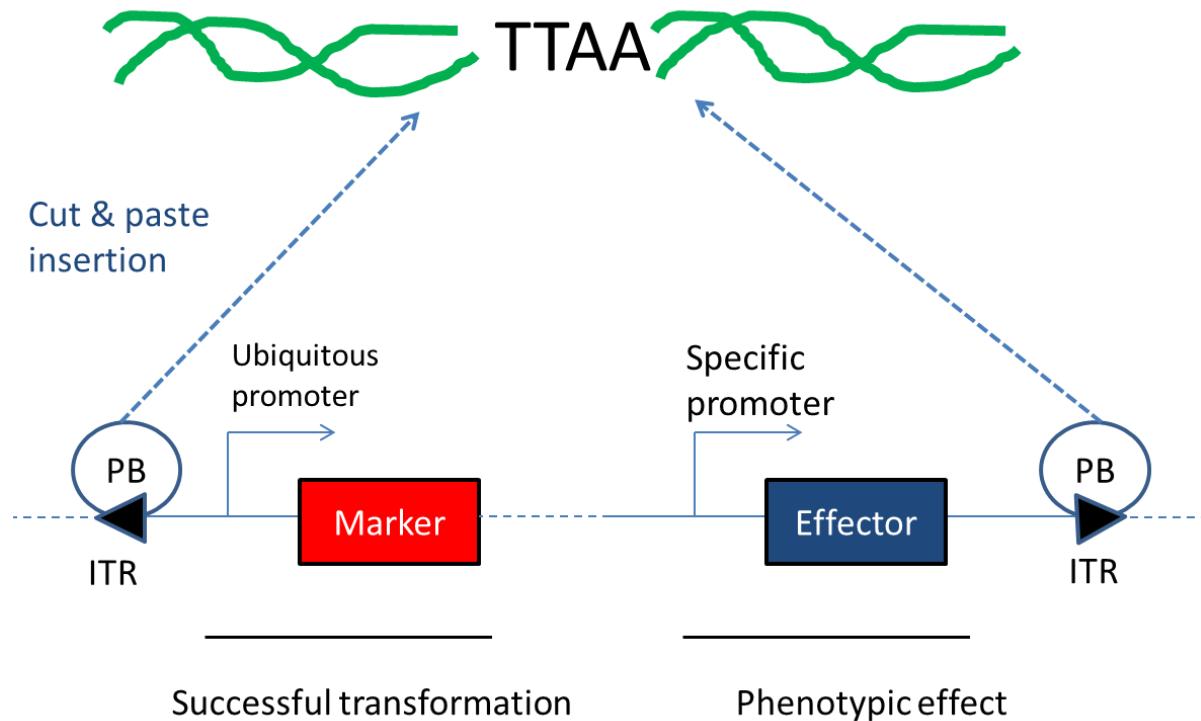


Figure 7. A typical *piggyBac* construct and its key genetic components (it is shown in a linear form here but is actually a circular plasmid). Successful transformation of the germline will be reported by a promoter and fluorescent marker combination. A tissue- and/or timing-specific promoter drive an effector to produce intended phenotypic effects such as female-killing. The transgenes are cut and pasted (as one) from the plasmid by the helper PB transposase which is injected in combination with the construct. The protein PB transposase recognises the transposon-specific inverted terminal repeat sequences (ITRs) located on both ends of the transposon vector (a circular plasmid) and moves the contents from the original sites and integrates them into TTAA chromosomal sites. The insertion site will vary from insertion to insertion and causes position effects or inter-line phenotypic variation.

1.1.11.1 PRIMARY MARKER AND PROMOTER INDICATIVE OF TRANSFORMATION

In transgenesis, expression of fluorescence is driven by a promoter and indicates successful insertion. A promoter is a recognition sequence upstream of a gene, a sequence to which RNA polymerase binds and around which the transcription initiation complex forms (Barrett *et al.*, 2012). The core promoter typically spans ~80 bp of the transcription start site (TSS). The downstream gene will be transcribed into mRNA, in this case a reporter fluorescent protein. The ideal promoter for this function should provide early, ubiquitous and strong expression for convenient screening.

The strength of a promoter depends on how closely it matches the ideal consensus sequence and resultant RNA polymerase binding rate. Consensus sequences for regulatory sites vary from species to species, and this is important with regard to exogenous gene expression in transgenic organisms (Levine & Tijan, 2003; Takahashi *et al.*, 2008; Clark & Pazdernik, 2012).

Visualisation of markers linked to these promoters can be impaired by insect morphology including wings and scales, or by autofluorescence of cuticle interfering with genuine fluorescence detection. In the Lepidoptera, recent work has characterised the function of three different regulatory sequences (*Opie2*, *Hr5ie1*, 3×P3) in driving expression of fluorescent proteins in diamondback moth (Martins *et al.*, 2012).

In some cases, strong tissue-specific promoters may be preferred, such as the artificial 3×P3, which has been used in *Tribolium* (Berghammer *et al.*, 1999, 2009; Martins *et al.*, 2012). Inclusion of the nuclear localisation signal also concentrates fluorescent protein in the nucleus, providing characteristic foci of fluorescence. This renders transgenic insects expressing fluorescent proteins easier to screen and distinguish from auto-fluorescent wild-type counterparts (Martins *et al.*, 2012). This is particularly conspicuous in cells with very large nuclei, for example those found in the Malpighian tubules.

1.1.11.2 TRANSGENE EFFECTOR EXPRESSION

In this thesis we are interested in two main classes of phenotypic expression and effect:

- i. Sperm-specific - as part of a sperm lethal system in diamondback moth.
- ii. Female-specific - as part of a genetic sexing system in *Tribolium castaneum*.

Success in engineering this tissue- or sex-specific expression is dependent on choosing appropriate endogenous genetic components to fulfil these requirements. These components regulate gene expression and are typically located in the untranslated gene regions and other non-coding elements (Figure 8).

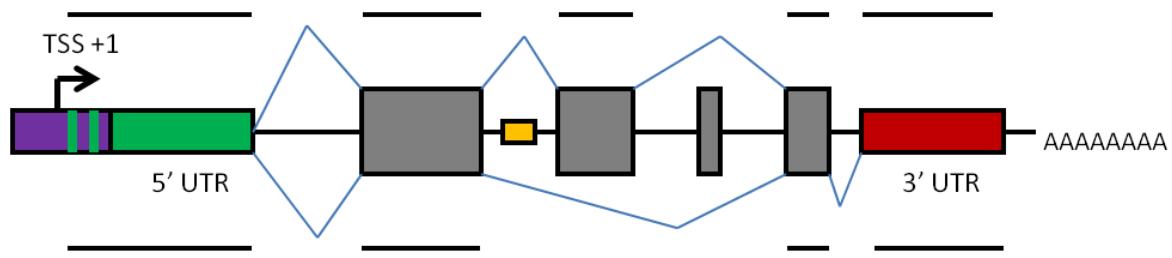


Figure 8. The standardised graphical representation of different genetic components of a construct adhered to for this thesis. Promoters are designated by arrows, untranslated regions (UTR) by thin rectangles, the relatively wide rectangles show exon positions with introns indicated using blue V-shaped lines, indicating the splice site. This one graphic displays two splice variants which in the context of this thesis we can call male and female. The exons are indicated in grey and orange rectangles for intronic enhancer elements. The same DNA sequence can provide different mRNA due to alternate splicing as shown by the different blue lines. The black lines represent the mRNA following splicing. The non-coding regions shown are important for gene regulation with complexity way beyond that of the simple molecular switches seen in prokaryotes (after Barrett *et al.*, 2012). The transcription start site, 5' and 3' untranslated regions are also shown. The poly-A tail is also shown at the 3' end, this may be involved in regulating translation and is distinguished by a distinctive poly-A signal in the 3' UTR.

Details pertaining to these general genetic components and their importance in construct design can be found in the appendix (Construct design general components and considerations).

1.2 AIM

This introduction has provided general background knowledge as a starting point. Each chapter has its own chapter-relevant introduction providing greater depth to the problem at hand and the relevant scientific background.

Specifically this introduction has described how this project will provide a knowledge-base for a paternal effect system in two model pest species, *Tribolium castaneum* and diamondback moth. The paternal effect system provides genetic control of the pest population by means of a repressible genetic system that can fragment male-gamete DNA in the release generation for AW- SIT. This will improve SIT in pest species where larval feeding causes economic damage. The developed genetic system should be compatible with a genetic sexing system. Both systems should be conditional (parallel or orthogonal). Ultimately the knowledge accrued should build

towards successful AW-SIT of the neglected lepidopteran and coleopteran pest species as part of an IPM framework.

1.3 OBJECTIVES

The thesis aims can be summarised as a series of objectives to strive towards. Each objective relates to a specific chapter. Each chapter will break these down into smaller chapter-specific objectives.

- Conduct an exploratory data analysis (EDA) of the transformation efficiency of *piggyBac* in insects to estimate parameters of key processes.
- Utilise these statistics to develop decision tools to manage expectations of probability of success and transformation efficiency in insect transgenesis using *piggyBac*.
- To engineer repressible male sterility phenotype in diamondback moth (“sperm-lethal system”).
- To engineer a transgenic strain of *Tribolium castaneum* using a fluorescent protein and endogenous promoter reproducing research from the literature.
- To develop a conditional female specific-lethal genetic system in *Tribolium castaneum*.
- To engineer repressible male sterility phenotype in *Tribolium castaneum* (“paternal effect system”).

The EDA provides knowledge to improve our understanding of the fundamental methodology associated with projects of this type that will be of lasting value to researchers in this field. The decision model developed from this EDA will optimise decision-making for insect transformation using *piggyBac*.

Working towards the sperm lethal system requires identification and validation of several distinct components (Figure 9). On starting this project both sexing and conditionality had been achieved in *Plutella xylostella*. In contrast *Tribolium castaneum* had none of the components in

place and was hitherto the first Coleoptera pest in the research group's laboratories. The story of the thesis can be displayed visually with our objectives for each species represented by a Venn diagram using set theory (Figure 9).

Chapter 2 relates to the exploratory data analysis of microinjection and transformation data in the literature and at Oxitec.

Chapter 3 uses the EDA to generate a hypothesis and test it experimentally as part of routine insect transformation.

Chapter 4 relates to the intersection of sets of sperm specific expression and conditionality in diamondback moth.

Chapter 5 relates to an intersection of different sets; sexing and conditionality in *Tribolium castaneum*.

Chapter 6 relates to an alternative approach to genetic sexing in *Tribolium castaneum* using RNAi instead of insect transgenesis.

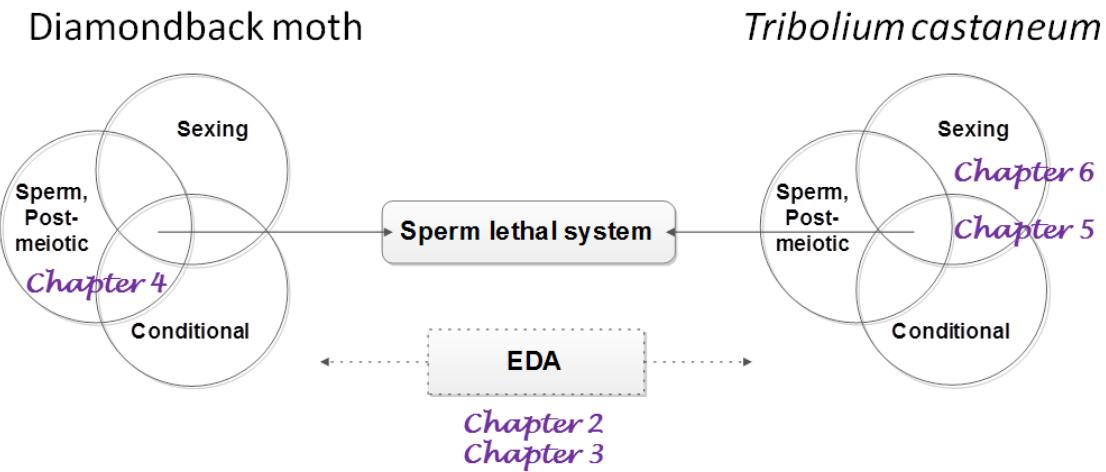


Figure 9. The story of this thesis – the paternal effect or sperm lethal system is dependent on prior success of transgenic insect modification with a desired phenotype including: a conditional or repressible system to facilitate mass rearing and genetic containment, a sexing or genetic sorting method for the sexes and precise sperm-specific, post-meiotic transcription and expression. Together these systems can be combined to create a sperm lethal system. The exploratory data analysis (EDA) provides a foundation for evidence-based decision making throughout focusing on the methods used to transform insects. The chapters are located in the corresponding Venn diagram intersects appropriate to that chapter.

CHAPTER 2: INSECT TRANSFORMATION WITH *PIGGYBAC*: GETTING THE NUMBER OF INJECTIONS JUST RIGHT

2.1. INTRODUCTION



2.1.1 INSECT TRANSFORMATION USING *PIGGYBAC*

The ability to integrate genetic constructs into the genome of organisms has utility in mitigating some of the global challenges facing humanity (Bazuin *et al.*, 2011; Kim & Pyykko, 2011; Morales *et al.*, 2007). Insect germline transformation (synonymous with insect transgenesis) can be employed to alter the phenotype of an insect by gene insertion (Fraser, 2012) and represents a research area attracting global interest (Tamura *et al.*, 2000; Handler

& Harrell, 2001; Perera *et al.*, 2002; Sarkar *et al.*, 2006; Morrison *et al.*, 2010; Raphael *et al.*, 2011).

Various methods can be employed to achieve genetic transformation. Transposable element (transposon) vector systems (Piegù *et al.*, 2015) were developed for *Drosophila melanogaster* using the *P* element (Rubin & Spradling, 1982). Though *P* works in only a very limited range of insect species, similar systems using other Class II transposable elements were developed for non-*Drosophila* insects. The most commonly used of these is the *piggyBac* element, originally discovered in cell lines of the cabbage looper moth, *Trichoplusia ni* (Fraser *et al.*, 1983; Sarkar *et al.*, 2003; Zimowska & Handler, 2006). It has been used for germline transformation in multiple insect orders (Table 3). A recent review by Kim & Pyykko (2011) summarises the molecular structure and mobility of *piggyBac*. The 2472-bp-long element is structured with two sets of inverted repeats at both ends and a central transposase-encoding open reading frame (Fraser, 2012). The insertion site of *piggyBac* is quasi-random, with a cut-and-paste insertion at the short genome motif site of TTAA (O’Brochta *et al.*, 2003; Wu & Burgess, 2004; Zhuang *et al.*, 2010).

For insect transformation, *piggyBac* constructs and the respective source of helper transposase are typically microinjected into pre-blastoderm embryos, with the offspring of the injection survivors examined for the expression of a marker gene, typically a fluorescent protein. A recent review describes the transposon vectors as having an “experimentally effective frequency, [however] the process remains relatively laborious and time consuming. Frequencies on the order of 0.1% to 10% are achievable, with higher frequencies less probable than lower ones” (Fraser, 2012).

Alternatives do exist including electroporation, ultrasonic activation and use of a “gene gun” (Wells, 2004; Mehier-Humbert & Guy, 2005; Al-Dosari & Gao, 2009). Other transposable

elements are used for insect germline transformation, and other molecular methods available, for example using Φ C31 or CRISPR/Cas9. This study is restricted to *piggyBac* as the most widely used method and correspondingly the one for which most data are available. This may provide a benchmark against which the efficiency of other methods may be compared.

Meta-analyses of data from multiple primary studies can be used to improve the efficiency of the scientific process (Brandt *et al.*, 2013) while simultaneously dispelling misconceptions (McClain *et al.*, 2015). Meta-analyses are usually associated with clinical trials and the medical literature, but recent co-opting of this technique has proven its applicability and usefulness to other scientific disciplines (Castellanos & Verdu, 2012). Meta-analyses facilitate the elucidation of effect sizes and inter-study variation despite noisy backgrounds associated with a typical single observational study.

Here we provide a description of the transformation efficiency of insect transgenesis using *piggyBac* as the vector. We draw upon a systematic literature analysis and an analysis of an unpublished data set provided by the biotechnology company, Oxitec Ltd. The application of the decision tools and information therein provides researchers with an approximation of what to expect when conducting insect transgenesis using *piggyBac*; complementing other attempts in the literature to characterise and quantify costs of genetic control (Alphey *et al.*, 2011).

Table 3. Summary of the earliest successful transformation of insect species using *piggyBac*. Modified from Morrison *et al.*, (2010).

| Family | Species name(s) | Reference |
|--|--|--|
| Mosquitoes | | |
| Culicidae | Yellow fever mosquito, <i>Aedes aegypti</i> | (Kokoza et al., 2001) |
| | Asian tiger mosquito, <i>Aedes albopictus</i> | (Labbé et al., 2010) |
| | <i>Aedes vexans</i> | (Rodrigues et al., 2006) |
| | New World malaria mosquito, <i>Anopheles albimanus</i> | (Perera et al., 2002) |
| | African malaria mosquito, <i>Anopheles gambiae</i> | (Grossman et al., 2001) |
| | Indo-Pakistan malaria mosquito, <i>Anopheles stephensi</i> | (Ito et al., 2002), (Nolan et al., 2002) |
| Fruit flies | | |
| Drosophilidae | Common fruit fly, <i>Drosophila melanogaster</i> | (Handler & Harrell, 1999) |
| | Spotted-wing drosophila, <i>Drosophila suzukii</i> | (Schetelig et al., 2013) |
| Tephritidae | Mexican fruit fly, <i>Anastrepha ludens</i> | (Condon et al., 2007b) |
| | Caribbean fruit fly, <i>Anastrepha suspensa</i> | (Handler and Harrell, 2001b) |
| | Oriental fruit fly, <i>Bactrocera dorsalis</i> | (Handler and McCombs, 2000) |
| | Queensland fruit fly, <i>Bactrocera tryoni</i> | (Raphael et al., 2010) |
| | Mediterranean fruit fly, <i>Ceratitis capitata</i> | (Handler et al., 1998) |
| Other Diptera (pest, myiasis, biting flies) | | |
| Muscidae | Housefly, <i>Musca domestica</i> | (Hediger et al., 2001) |
| Calliphoridae | Australian sheep blowfly, <i>Lucilia cuprina</i> | (Heinrich et al., 2002) |

| | | |
|------------------------------|---|---------------------------|
| | New World screwworm, <i>Cochliomyia hominivorax</i> | (Allen et al., 2004) |
| Diopsidae | Stalk-eyed flies, <i>Teleopsis dalmanni</i> | (Warren et al., 2010) |
| Wasps, bees and ants | | |
| Hymenoptera | Sawfly, <i>Athalia rosae</i> | (Sumitani et al., 2003) |
| | Honeybee, <i>Apis mellifera</i> | Schulte et al., 2014 |
| Beetles | | |
| Coccinellidae | Harlequin ladybird, <i>Harmonia axyridis</i> | (Kuwayama et al., 2006) |
| Tenebrionidae | Red flour beetle, <i>Tribolium castaneum</i> | (Berghammer et al., 1999) |
| Butterflies and moths | | |
| Nymphalidae | Squinting bush brown butterfly, <i>Bicyclus anynana</i> | (Marcus et al., 2004) |
| Gelechiidae | Pink bollworm, <i>Pectinophora gossypiella</i> | (Peloquin et al., 2000) |
| Bombycidae | Silkworm, <i>Bombyx mori</i> | (Tamura et al., 2000) |
| Plutellidae | Diamondback moth, <i>Plutella xylostella</i> | (Martins et al., 2012) |
| Crambidae | Asian corn borer, <i>Ostrinia furnacalis</i> | (Liu et al., 2012) |
| Tortricidae | Codling moth, <i>Cydia pomonella</i> | (Ferguson et al., 2011) |

2.2. RESULTS AND DISCUSSION

2.2.1 SYSTEMATIC REVIEW OF TRANSFORMATION EFFICIENCIES IN PUBLISHED LITERATURE

Design and implementation of meta-analysis followed guidelines in Khoshdel *et al.*, 2006.

Meta-analyses tend to be conducted in the medical literature, so methods were co-opted as appropriate (Reade *et al.*, 2008; Cooper and Patall, 2009). The structure of the methods section follows Sim *et al.* (2011). A checklist for evaluation of meta-analysis quality is described by Huf *et al.* (2011). Full details can be found in the supplementary material (Appendix 9.7.1).

2.2.1.1 DATA SOURCES

A summary of the literature search can be seen in Figure 10. The following electronic databases were searched from inception to March 2013, repeated in October 2013 and March 2015, to identify relevant experiments and or studies: Web of Knowledge, PubMed and Scopus databases. The key terms used for the search were; ([*piggyBac* AND insect*]). The database search results were refined by manual inspection and identification of publications with relevant transformation efficiencies. The title and then abstract were read. The following checklist was applied to candidate studies to be included in the analysis:

1. The species transformed is an insect.
2. The insect germline was transformed using the *piggyBac* vector.
3. The *piggyBac* vector was microinjected into embryos.
4. Injection data are included: specifically number of injected embryos, number of injection survivors and number of independent transgenic lines generated per unique

construct (the derived variables; survival and transformation efficiency, were calculated from the raw data).

Typical reasons for non-inclusion were: different methods of transgenesis; transgenesis in cell lines rather than the whole organism (for example in Mandrioli & Wimmer, 2003) and inter-database duplication. Following this process 32 studies remained (Figure 10; Table 3). Additional details concerning the data extraction methods, summary statistic of choice and bias considerations are given in the supplementary material (Appendix 9.7.1).

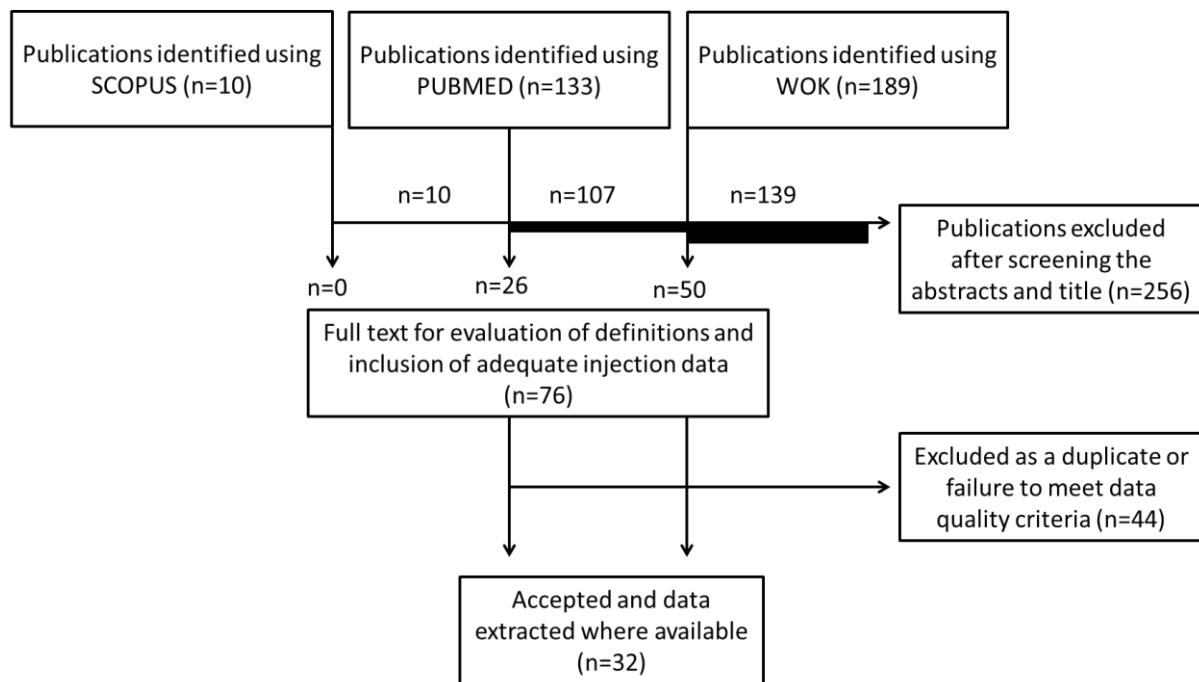


Figure 10. Flow chart of the paper selection process using the search term, across three life science relevant databases (SCOPUS, PUBMED and Web of Knowledge (WOK)). 32 publications describing 86 unique experiments provided microinjection data and transformation efficiency estimates across a range of insect species.

2.2.2 TRANSFORMATION EFFICIENCY BY INSECT ORDER AND SPECIES FROM PUBLISHED DATA

Germline transformation or transgenesis has been achieved across a diverse range of insect Orders (Table 3). Some authors have hinted at a difference between transformation efficiencies between Orders, with the Lepidoptera efficiencies being lower compared to the

Diptera for example (Marec *et al.*, 2005). The data were plotted to examine this, at Order level (Figure 11A) and at Species level (Figure 11B). Most of the transformation efficiency estimates (52/74) were from transformed dipteran species, with 22 of those of the genus *Drosophila*. All medians were between 0.001 and 0.1 except for the coleopteran estimate of 0.237 (see Appendix 9.7.1.2 for discussion of appropriate statistics and methods to describe the distributions). The lower whisker in the Coleoptera is the data point provided by the only non-*Tribolium castaneum* transformed beetle; the ladybird, *Harmonia axyridis* at 0.0370 (Appendix 9.7.1.3 for Bayesian methods to produce a posterior probability distribution for the transformation efficiency of a species). The Hymenoptera have only one representative so should be excluded from comparison.

Outliers are a common characteristic in each Order with some transformation efficiencies of over 0.3 occurring. Following inspection the outliers are produced by less precise transformation efficiency mean estimates due to a relatively small number of trials for that particular *piggyBac* insect combination. For example, the lepidopteran *Bombyx mori* produced an outlier experiment with five transformed lines from 27 G₀. As pointed out by Fraser (2012), the more extreme the transformation efficiencies the less frequently those efficiencies are observed.

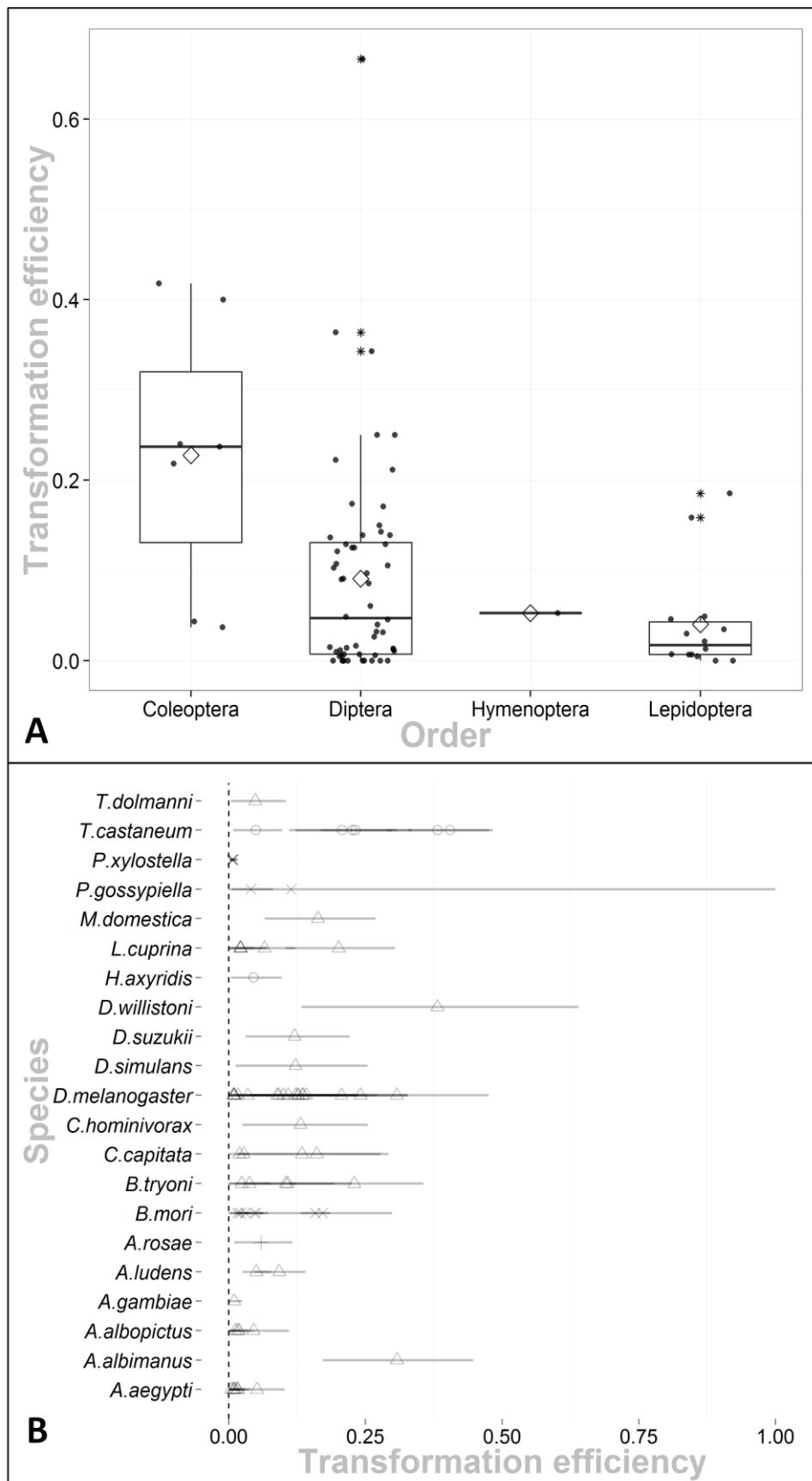


Figure 11. The transformation efficiency of *piggyBac* in insects. A) A box-and-whisker plot of the meta-analysis transformation efficiency data sub-setted by insect order. Each sub-set was comprised

of, n equal to: Coleoptera, seven; Diptera, 52; Hymenoptera, one; and Lepidoptera, 14. The dark line represents the median, the box the interquartile range (IQR) and the whiskers highest value within 1.5 times the IQR above quartile 0.75 (or equivalently below quartile 0.25) in a standard boxplot, the black asterisks are supplementary to the scatterplot and identify horizontally adjacent outliers within an order. The Hymenoptera and Lepidoptera do not have whiskers plotted as all non-anomalous data are found within the IQR. A horizontal jitter plot is superimposed onto the boxplot showing the transformation efficiency of each unique construct species combination found within the literature search. The grand mean by insect order is represented by the empty diamond. B) Published transformation efficiencies in insects found by this systematic analysis, sorted by species (alphabetical order by species). The order to which the species belongs is represented by the shape of the points (Diptera; triangle, Coleoptera; circle, Lepidoptera; cross and Hymenoptera; plus). Mean estimate from individual experiments shown by the shapes. Horizontal lines represent the upper and lower 95% confidence intervals of each experimental mean calculated using Bayesian methods (with prior distribution provided by beta distribution fitted to the combined data; shape 1 = 0.73 and shape 2 = 5.67) in R with the package "binom" (Dorai-Raj, 2014). Each experimental construct species combination has its own mean and confidence interval, the transparency of the points and intervals allows overlap to be visualised.

2.2.3 ANALYSIS OF AN EXTENSIVE UNPUBLISHED DATASET OF PIGGYBAC TRANSGENESIS EXPERIMENTS

The biotechnology company Oxitec has collected a dataset of over 250,000 insect injection experiments (Figure 12A) using *piggyBac* (Table 4). The data were collated and subjected to exploratory data analysis of the derived variables microinjection survival and transformation efficiency to establish typical values and any discrepancy from the published data set.

Table 4. Summary of the dataset; sub-sets organised by species. Information includes number of unique constructs injected into a species (the evaluation unit), sum of embryos injected (n_1), sum of injection survivors to fertile adults (n_2 or G_0), and total number of independent transgenic lines created in that species in the compiled Oxitec dataset. Accurate as of March 2014.

| Species | Sum of unique constructs | Sum of embryos microinjected | Sum of microinjection survivors | Sum of independent transgenic lines |
|---------------------------------|--------------------------|------------------------------|---------------------------------|-------------------------------------|
| <i>Aedes aegypti</i> | 46 | 71, 252 | 3, 314 | 239 |
| <i>Aedes albopictus</i> | 10 | 37, 235 | 5, 339 | 89 |
| <i>Ceratitis capitata</i> | 26 | 21, 858 | 5, 977 | 167 |
| <i>Bactrocera oleae</i> | 5 | 27, 500 | 760 | 23 |
| <i>Drosophila suzukii</i> | 4 | 3, 287 | 138 | 4 |
| <i>Pectinophora gossypiella</i> | 37 | 55, 605 | 9, 296 | 100 |
| <i>Plutella xylostella</i> | 34 | 68, 547 | 21, 761 | 108 |
| <i>Tribolium castaneum</i> | 2 | 5, 227 | 572 | 8 |
| <i>Tuta absoluta</i> | 2 | 7, 244 | 601 | 3 |

2.2.3.1 DATA ENTRY AND CHECKING

A rectangular data set was compiled using data accrued from more than six years of research involving the microinjection of nine insect species' embryos with exogenous DNA. The data was transliterated from the original laboratory books as well as student theses (Bilski, 2012; Ant, 2013; Harvey-Samuel, 2014), by personal communication, and where published cross-checked against publications. Data were organised in a “tidy” data frame (Wickham, 2014) and validated by re-entry. Missing data were treated as described in the Supplementary Material (9.7.1.4.2). The full data set is available from <https://github.com/mammykins/piggyBac-data>.

2.2.3.2 TRANSFORMATION EFFICIENCY DISTRIBUTION

A summary of the data is provided by plotting the injection survival to adulthood against the transformation efficiency for each evaluation unit (Figure 12B) by insect order. This facilitates comparison to the meta-analysis as well as Fraser's description of the *piggyBac* transformation efficiency interpretation (Fraser, 2012). This confirms the comments of Fraser with most of the data distributed between 0.01 and 0.1 (117/166 or 70% observations lay within this range).

Interestingly we see a clustering among the orders with the Lepidoptera tending to have lower transformation efficiencies (0-0.1) compared to the Diptera, where it is not uncommon to have transformation efficiencies above 0.1. This does not necessarily mean Lepidoptera are more difficult to transform, as the Lepidoptera injected embryos are more likely to survive. This may be accounted for by differences in injection methodology or by the hardiness of the embryo.

This supports previous work that has compared the variability between efficiencies in the Diptera and Lepidoptera. Lobo *et al.*, (1999) compared the mobility of *piggyBac* in embryos from different insect families using a transposition assay. The rate of transposition in dipteran species was higher than that of *T. ni*, which harbours the *piggyBac* transposon (Mohammed & Coates, 2004).

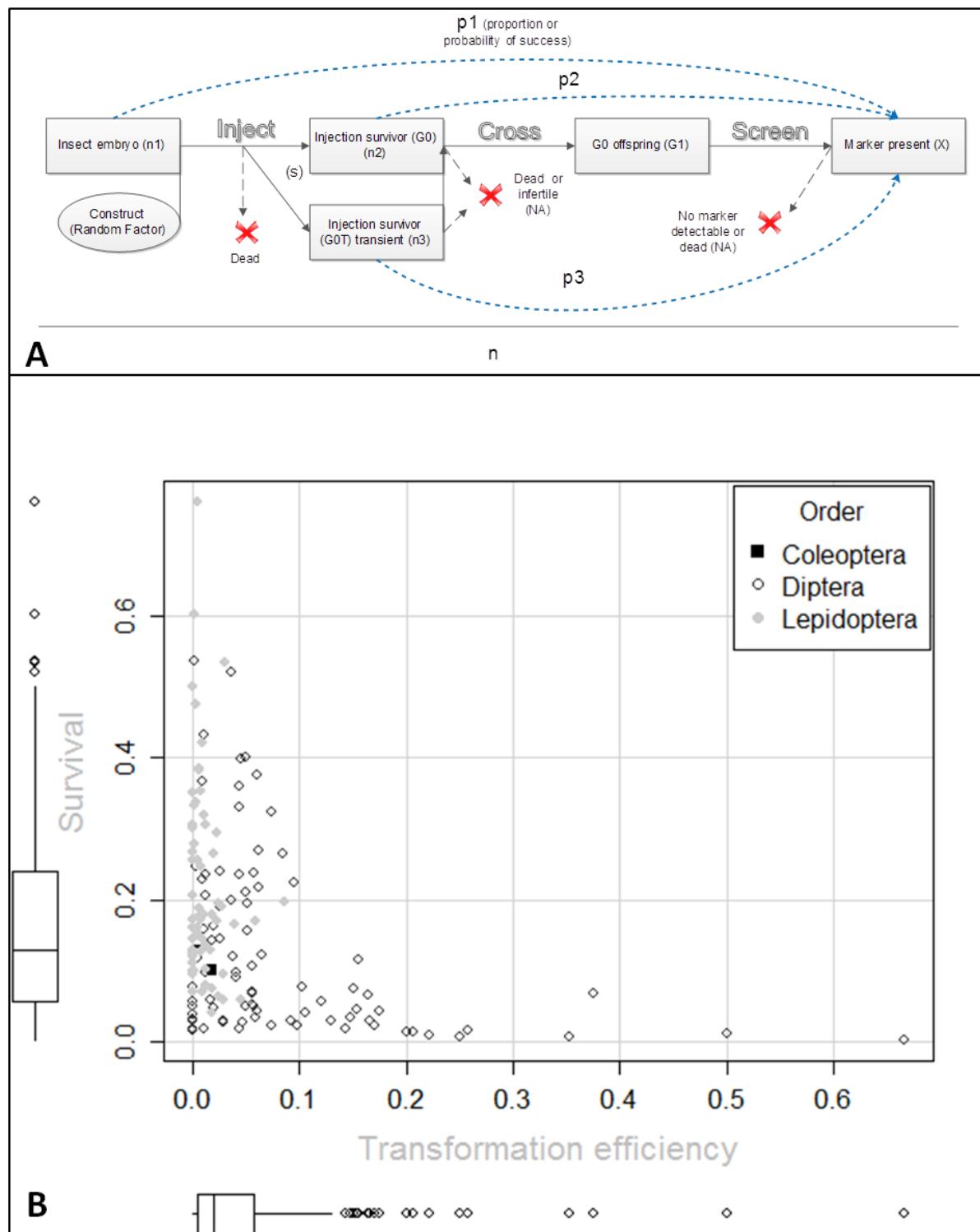


Figure 12. A) The experimental unit set up of variables and statistics of interest after a given construct was injected into an embryo, n_1 times, with probability of survival s . Values not observed or recorded are annotated with “NA”, not available. The outcome of each trial, n , is assumed to be independent from the outcome of all other trials. After injection the embryo has p_1 probability of producing transgenic offspring. The embryo will either die, or develop into a transient (G_0T) (injection survivors showing transient expression of the fluorescent marker) or non-transient fertile adult (G_0) (injection survivor not showing transient expression of the fluorescent marker) (or a transient adult that was not detected as transient). Considering the G_0 survivors, n_2 – a proportion, p_2 will produce at least one transgenic offspring (known as “transformation efficiency” in Warren *et al.*, 2010 and Martins *et al.*, 2012). This particular transformation efficiency (X/G_0 and X/n_2) leads to a lower calculated rate than if only fertile G_0 individuals were considered;

as we have no estimate of the infertility or fertility rate in G_0 ; this could not be corrected for. Multiple transgenic G_1 from the same G_0 parent pool are assumed to represent a single transformation event unless shown otherwise with molecular tests. Of the G_0T , n_3 – a proportion, p_3 will produce transgenic offspring. The proportions p_1-p_3 are bounded between zero and one and are derived from how many times an event, transgenesis (X), did or did not occur (the numerator). The phrases transgenic efficiency and transgenic rate are used interchangeably for p_2 (p_1 is not given a name despite recommendations from Warren et al., 2010). B) The correlation between microinjection survival and the achieved transformation efficiency is described. Each point represents an evaluation unit, the number of successes of a unique construct injected into a species of pre-blastoderm embryo, divided by the number of trials. The survival axis corresponds to number of injections survivors (to adulthood) divided by the number of embryos injected. The transformation efficiency is derived from the number of injection survivors divided by the number of independent transgenic lines generated. The figure does not group the data into species hence the more uniform and poly-modal shape along the survival axis. A box-and-whisker plot on each axis describes the density of the data.

2.2.3.3 PUBLICATION BIAS

For those species for which we have approximately 30 or more experiments with transformation efficiency data, we plotted all the data and examined those that have been published (see Appendix 9.7.1.1.3). We observe many more zero and near zero transformation efficiencies than would be expected given the meta-analysis findings hinting at bias (see Appendix 9.7.1.1.4 for quantification of the bias). We also observe that atypically-high transformation efficiencies tend to be associated with a lower number of injection survivors. This could be caused by researchers stopping their inspection and screening of G₀ crosses when they feel they have enough lines generated thereby overestimating the efficiency. Conversely the lower efficiencies at the higher number of injection survivors crossed could be caused by researchers not stopping until they have success. It is also unclear whether all zero successes experiments are recorded.

2.2.3.4 INTER-SPECIES VARIATION IN SURVIVAL

The distributions were visualised using a boxplot and scatterplot hybrid (Figure 13A). The distributions located further away from the bounds (zero and one) tend to be less skewed and more variable. Extreme values near one or zero are improbable, indeed if zero survival were achieved the results may have been discarded; furthermore, one hundred percent survival does not occur even with un-injected embryos in optimal conditions. Seventeen of the 166 experiments did not include survival data due to the number of injections or injection survivors missing for the experiment in question. There were no recorded experiments with a zero survival. The maximum survival was achieved in *Plutella xylostella* with 0.76 compared to the lowest non-zero survival of *Aedes aegypti* at 0.0028.

2.2.3.5 INTER-SPECIES VARIATION IN TRANSFORMATION EFFICIENCY

Most of the data are found between one and ten percent (Figure 13B). However, some species appear to be highly clustered, with all *P. xylostella* data found between zero and five percent inclusive. As pointed out by Fraser (2012) the more extreme the transformation efficiencies the less frequently those efficiencies are observed. This can be envisioned as a long tail or a skewed positive distribution. This is observed with the efficiencies far away from the main cluster, as seen in the Diptera and Lepidoptera. For those species with greater than ten data the interquartile range (IQR) tends to increase as the median moves away from zero. The tails of both mosquito distributions extends above 20%, with outliers for *Ae. aegypti* as high as 66% (the maximum achieved). Closer inspection reveals the datum responsible comprised of two transgenic lines from three G₀. This species has been described in an earlier study as having a typical transformation efficiency of only 8% (Nimmo *et al.*, 2006).

The Lepidoptera have consistently lower transformation efficiency compared to Diptera. However, both orders have their share of zero transformation efficiency experiments with 27 in total between them. Those species with lower median transformation efficiency have more zero transformation efficiency experiments. The two Lepidoptera, *P. xylostella* (9/31) and *Pectinophora gossypiella* (9/35), have nine each compared to one *Ceratitis capitata* (1 /26) and five *Ae. aegypti* (5/39) zero experiments. 19.4% per cent of experiments (27/139) have ended without germline transformation contrasted to the meta-analysis literature rate of 12% (9/75). This could suggest that those species with lower transformation efficiencies are more likely to have a construct abandoned, or microinjection of DNA is mechanically more difficult, with more injection survivors not containing any plasmid DNA, or simply due to natural variation in the insertion rate.

Despite the nature of the inter-experimental variation, such as different constructs injected, different engineers and rearing methods, the transformation efficiency is remarkably consistent within some species, particularly the Lepidoptera. This may partly arise as a relic due to the enforced bounding at zero, however it does suggest that given a new construct, it is possible to provide a reliable prediction of the transformation efficiency.

A caveat, some constructs encoded dominant lethal genes as part of a repressible system, and so one might expect a proportion of transgenics to be lost due to transient (episomal) expression. We might expect this to produce a bi-modal distribution of transformation efficiency in a species. This may be evident in *Aedes aegypti* (Figure 13B).

2.2.3.6 INTER-SPECIES VARIATION BY CONSTRUCT SIZE

Experimental evidence for other transposons suggests a negative correlation between the size of the construct and the transgenic efficiency of the vector (Delattre *et al.*, 2000). This could be a consequence of larger DNA molecules diffusing smaller distances (Lukacs *et al.*, 2000), larger plasmids being injected at a lower molar concentration relative to smaller plasmids or a physical limitation of the vector. Typically a construct injected will be between 10,000 and 15,000 bp in length. It is therefore not recommended to extrapolate the data and try to identify a trend where very few values lie outside the typical range. This is compounded in some species by the low number of constructs injected; each point has a large effect on the overall trend, where removal or addition of one datum can change the inference made (Figure 13C).

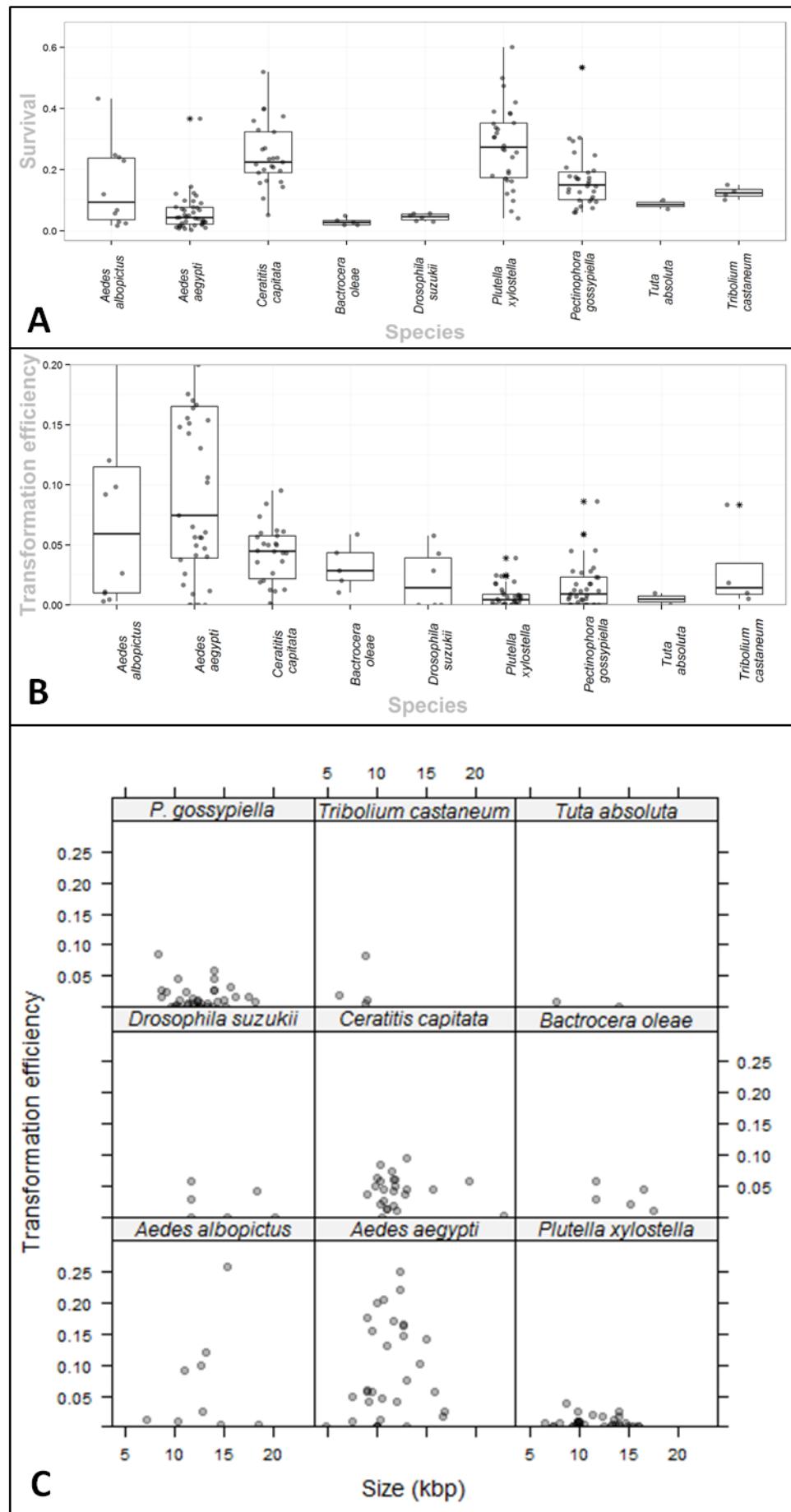


Figure 13. A) The proportion survival of Oxitec insect research species, from embryo to fertile adult, following microinjection of a *piggyBac* vector. A horizontal jittered scatterplot is overlayed on a boxplot, summarising the survival distribution for each species. Species are grouped by insect order (Diptera, Lepidoptera and Coleoptera) from left to right. The asterisks show the points that were outliers. For each species there are 10 *Aedes albopictus*, 39 *Aedes aegypti*, 26 *Ceratitis capitata*, five *Bactrocera oleae*, four *Drosophila suzukii*, 29 *Plutella xylostella*, 32 *Pectinophora gossypiella*, two *Tuta absoluta*, and two *Tribolium castaneum* data, each representing a unique construct experiment. One outlier at 0.67 survival was removed from *Plutella xylostella* to improve ease of reading. B) The transformation efficiency of different genetic constructs vectored by *piggyBac* into the germline of different insect species. Transformation efficiency is defined as the number of independent transgenic lines divided by the number of fertile injection survivors crossed, given the unique construct species combination. Few data were found above 0.2 transformation efficiency so the y axis was limited to this range. For each species there are 10 *Aedes albopictus*, 39 *Aedes aegypti*, 26 *Ceratitis capitata*, 5 *Bactrocera oleae*, four *Drosophila suzukii*, 31 *Plutella xylostella*, 35 *Pectinophora gossypiella*, two *Tuta absoluta* and two *Tribolium castaneum* data. C) A lattice plot of the transformation efficiency of constructs injected into different species by size (in base pairs). Most constructs are 10,000-15,000 bp in length. As these are injected more frequently there is a cluster of points around this range for each species. Small (< 10 kbp) or large (> 15 kbp) are injected more rarely and the data is sparse. The data is subsetted into a species pane with the species label above. Each datum is transparent, dark points represent overlap.

2.2.4 GOLDILOCKS DECISION-MAKING: HOW TO GET THE NUMBER OF INJECTIONS JUST RIGHT WITH *PIGGYBAC*

Injecting too few embryos can result in no or very few transgenic lines. If the lines generated do not show the desired phenotype the investigator is left uncertain as to whether the construct needs to be redesigned or it failed due to position effects. The other extreme involves an excessive number of injections, as the investigator urgently seeks to generate at least one transformed line. Historical data provide an opportunity to estimate the transformation efficiency and guide future experiments.

2.2.7.1 AN EXAMPLE USING PLUTELLA XYLOSTELLA DATA

The *P. xylostella* survival and transgenesis efficiency are highly skewed rendering the mean a poor descriptor of the central location of either distribution (Figure 14B). The median is a better metric as it is more robust to extreme values and because several (9/34) zero values were also present. A representation of the black box model is shown in Figure 14A. The probability of the embryo failing to achieve G₀ status is 1-s (where s is the median survival of an embryo to G₀ post-injection). The embryo survives injection, hatches and the larva develops to adulthood with probability s. The G₀ (assumed fertile adult) is then crossed and the offspring G₁ are screened for the transgene. A G₀ gives rise to a unique insertion event with probability X / G₀ (where X is the number of independent lines produced). Published and Oxitec data give the probability of transformation of *P. xylostella* under this model as 0.0065 and 0.0043, respectively (Appendix 9.7.1.4.8).

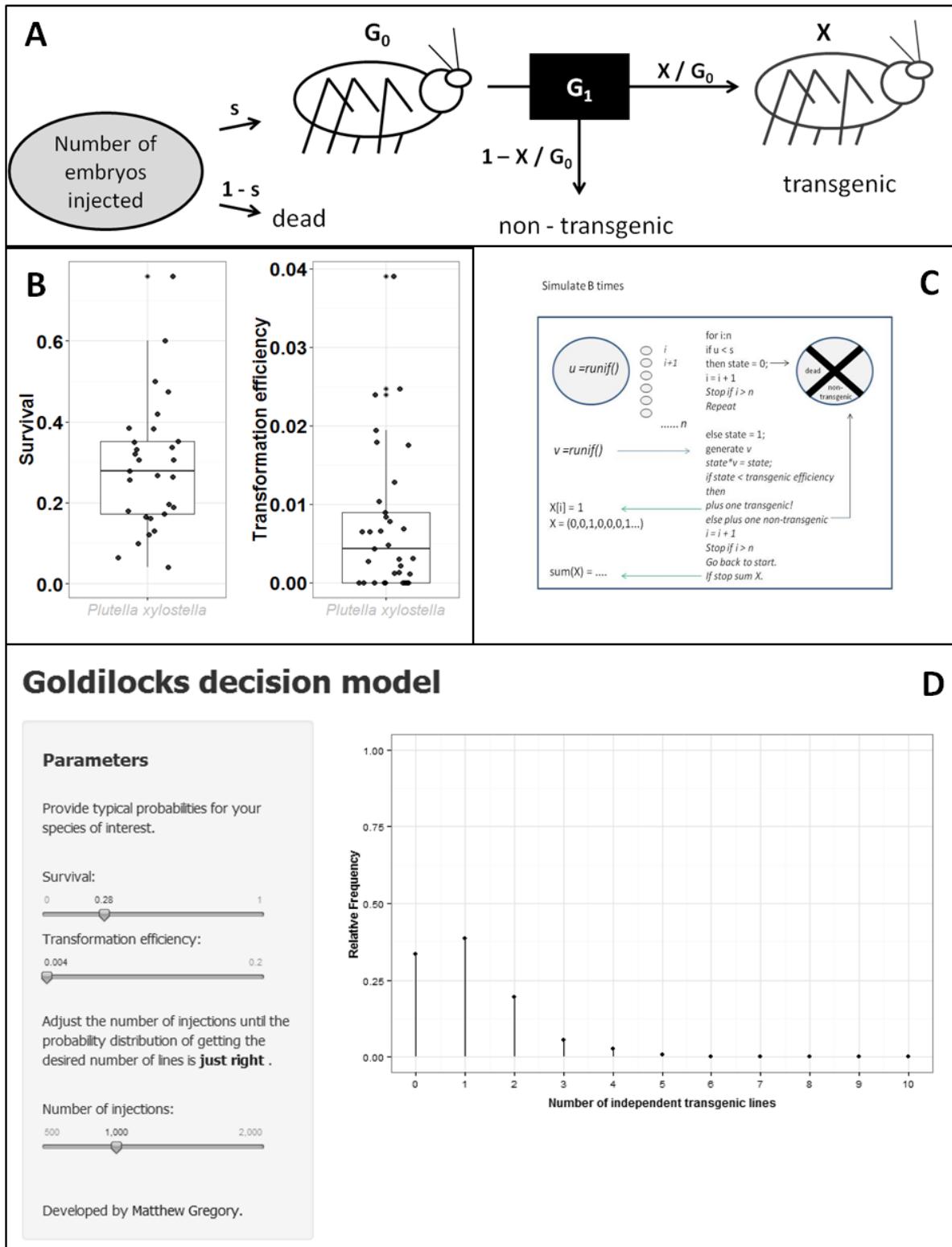


Figure 14. A) A simple black box model depicting the trajectory of one embryo injected with a novel construct never before injected. The probability the embryo survives injection, s , is based on previous injection data, as is the transformation efficiency, X/G_0 . For a given number of independent injections, n , the number of transgenic lines can be estimated given that particular simulation of the model. The stochasticity at each branch adds variability to the output. Accordingly the simulation for n injections should be repeated an appropriate number of times to provide a discrete probability distribution of the frequency of transgenic lines produced. B) A jittered box and whisker plot of the survival ($n = 29$) and transformation efficiency ($n = 33$) distributions in diamondback moth (*Plutella xylostella*). The median values are 0.28 and 0.0043, respectively.

Outliers are annotated with an adjacent asterisk. C) A diagram describing the approximate structure of the program used to model the number of independent transgenic lines produced from n injections over B simulations. The pseudocode is simplified and does not match precisely how the MATLAB / R function works. The reader is advised to start from the top and read from left to right. The variables “s” and “transformation efficiency” are input into the model beforehand and are fixed following the example of using the median survival and transformation efficiency. The branching steps provide the stochasticity of the model whereby u and v are drawn from a random uniform distribution between 0 and 1. These numbers are compared with the input variables thus determining the fate of the injected embryo. At the first branch it is determined whether the embryo survives (it remains “alive”, coded as 1) and at the second whether it gives rise to at least one transgenic offspring. State is a placeholder variable which deals with the previous logic branch by converting the alive embryo to the current value of v (multiplying by v, if alive state equals v, if dead state equals zero). At the end of the loop, the process is repeated, unless the desired injection number has been reached. The number of transgenics is recorded in a vector X and summed upon reaching n injects. This provides an integer which again is stored in a vector of B length. This vector provides the information to draw a discrete probability distribution of the expected number of transgenic lines produced from n injections simulated B times. D) The Goldilocks application interface for helping researchers get the number of injections just right. The output updates when the slider inputs are changed. The most recent data for *Plutella xylostella* are used.

2.2.7.4 A DECISION-MAKING MODEL FOR INSECT TRANSGENESIS

A Markov-Chain Monte-Carlo simulation (pseudocode in Figure 14C) was used to model the system where the number of embryos injected, n , gives a binary vector of successes or failures. The final state of each embryo either gives rise to a transgenic-bearing G_0 (1) or it does not (0; dead or non-transgenic). Stochasticity is built into the model at each branch where a pseudorandom number (between 0 and 1) is generated and tested against the input parameters as appropriate. The simulation can be run repeatedly to estimate the discrete probability distribution of the total number of transgenic lines produced by n injections. The model was initially developed in MATLAB 2012a Student Version (The MathWorks, Inc., Natick, Massachusetts, United States) then re-coded in R using R studio (<http://www.rstudio.com/>) and R shiny (<http://www.rstudio.com/products/shiny/>) to develop a web application for insect transgenesis researchers to use. The model is named “Goldilocks”, to assist researchers in getting the number of injections just right (Figure 14D). The model is available online at <https://mammykins.shinyapps.io/App-gold> and can be implemented locally by using the code from <https://github.com/mammykins/Goldilocks-decision-tool>. A graphic interface allows the user to adjust sliders to the appropriate values for a species of interest (suggested values are provided in Table 5).

Table 5. The median survival and transformation efficiency achieved in species transformed at Oxitec. The statistics in bold are based on many experiments and are likely more reliable. The lowest transformation efficiencies belong to the Lepidoptera.

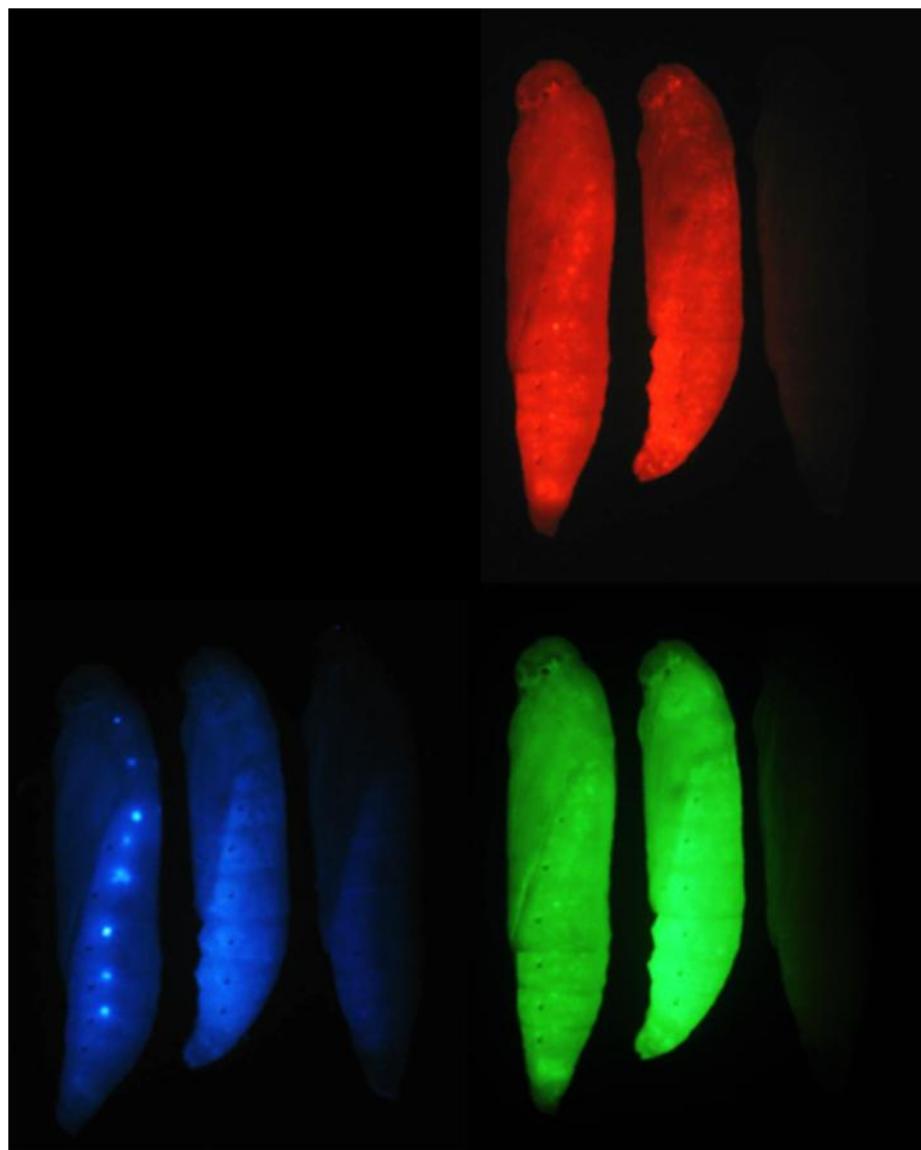
| Species | Order | Survival | Transformation efficiency |
|---------------------------------|-------------|--------------|---------------------------|
| <i>Aedes aegypti</i> | Diptera | 0.093 | 0.059 |
| <i>Aedes albopictus</i> | Diptera | 0.042 | 0.074 |
| <i>Bactrocera oleae</i> | Diptera | 0.028 | 0.029 |
| <i>Ceratitis capitata</i> | Diptera | 0.230 | 0.045 |
| <i>Drosophila suzukii</i> | Diptera | 0.046 | 0.014 |
| <i>Pectinophora gossypiella</i> | Lepidoptera | 0.149 | 0.009 |
| <i>Plutella xylostella</i> | Lepidoptera | 0.278 | 0.004 |
| <i>Tribolium castaneum</i> | Coleoptera | 0.124 | 0.014 |
| <i>Tuta absoluta</i> | Lepidoptera | 0.086 | 0.004 |

2.3. CONCLUSION

This paper provides a rigorous description of the distribution of microinjection survival and *piggyBac* transformation efficiencies in different insect species by inspection of the published literature and analysis of an experimental data set. This allows insight into publication bias and misconceptions of what is a typical survival or transformation efficiency in a given species. Combined with the Goldilocks decision model, researchers can use this analysis to minimise wasted effort and resources due to an inappropriate number of injections being carried out. This tool can also be applied to other transposons or transgenesis methodologies in any species provided the survival and transformation efficiency statistics are available.

In the future, as our dataset increases, we should look to more sophisticated machine learning techniques to identify patterns in the data. These tools, such as logistic regression, could provide a predicted transformation efficiency given the specific characteristics of the construct-species combination.

CHAPTER 3: WHAT COMPONENTS OF A MICROINJECTION MIX ARE REQUIRED FOR TRANSIENT EXPRESSION OF A MARKER?



3.1 INTRODUCTION

Chapter 2 discussed how an EDA could be used to generate hypotheses to be tested during routine research generating transgenic lines. This chapter represents an implementation of that advice.

The motivation for the experiment was based on mixed researcher opinion (Alphey, Morrison, Koukidou, Walker, Harvey-Samuel; personal communication). Consequently I set out to design an experiment which focuses on a niche area of insect transgenesis.

Is transient or episomal (Morange, 2009) expression of a fluorescent protein marker — expression in the somatic tissue that has not yet been confirmed as heritable — more likely to arise in those injected embryos where both the construct plasmid and the helper transposase are present (Figure 15)? In this thesis the adjective transience is defined as, “showing transient-expression of a reporter gene”. A transient, a noun, is defined as an organism that is positive for transience.

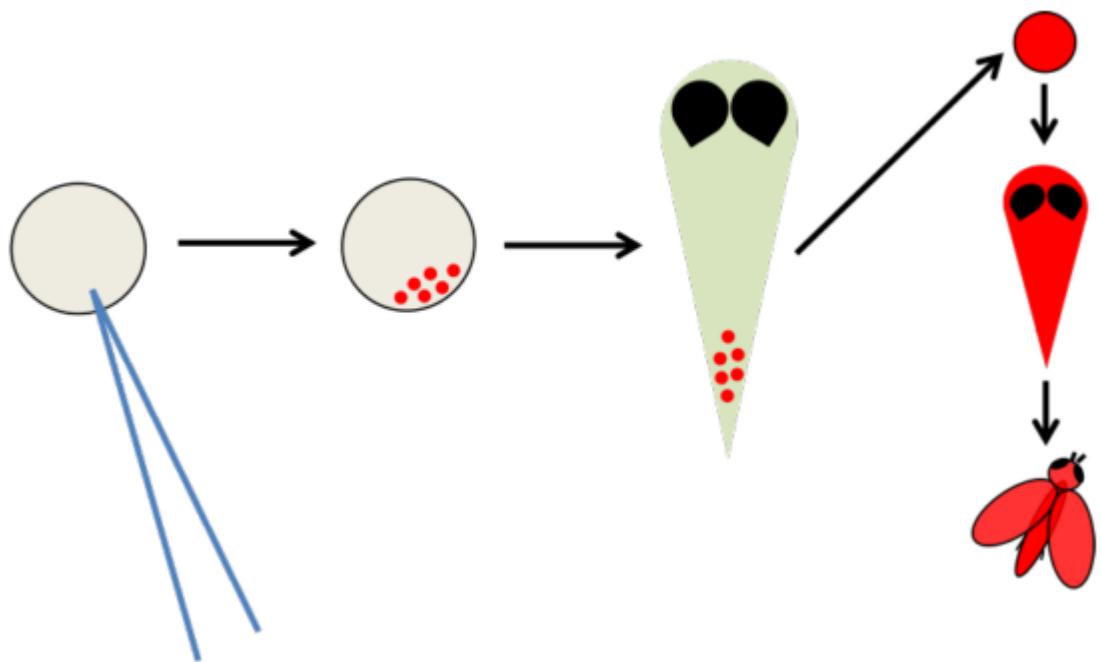


Figure 15. Schematic showing the process of transformation and the position of transient fluorescence in this process. From left to right, (1) eggs are microinjected at the posterior pole. (2) Transgenes may begin to express in the developing embryo and are detectable if coding for fluorescent proteins. This could be a result of either chimeric expression (where the transgene has integrated into individual cells) or through episomal expression (transcription and translation from injected construct plasmid without integration). (3) Transient-expression, where present in embryos, may or may not be visible in developing pupae. If integration of the transgene has occurred into germ-line cell(s), gametes produced by this individual as an adult may contain a genomic copy of the transgene. If these gametes successfully contribute towards a zygote (F1 generation (4)), and this zygote is identified, a new transgenic line has been identified (after Harvey-Samuel, 2014).

The common sense view is that the transience is a proxy indicating that both the construct and transposase have been successfully injected into an embryo. The belief extends to the notion that those injection survivors that display transience are more likely to give rise to offspring with carrying an insertion of the *piggyBac*-vectored transgene. By extension, some research groups assign special treatment to these insects (by using them in more favourable cross ratios) or in extreme cases discarding all non-transient injection survivors (Alphey, personal communication).

At Oxitec, during the historic generation of transgenic lines, the Lepidoptera were crossed under the following conditions:

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

- If non-transient, the male was crossed with 10 non-transient females plus 10 wild-type females.
- If transient, the male was crossed with three wild-type females and vice versa.

This data was explored to detect any correlation between transience and transformation efficiency.

3.2 EXPLORATORY DATA ANALYSIS

The transformation efficiency in the Lepidoptera has been shown to be relatively low at typically less than 1%. Plotting transformation efficiency on a log axis reveals how near zero the efficiency can go (Figure 16). If we compare transience rate to transformation efficiency we may expect a positive correlation as transience confirms a successful injection of construct DNA into an embryo. However, as all embryos are injected, perhaps transience is a superfluous proxy, as we would expect both construct and helper to be present in the embryo given injection.

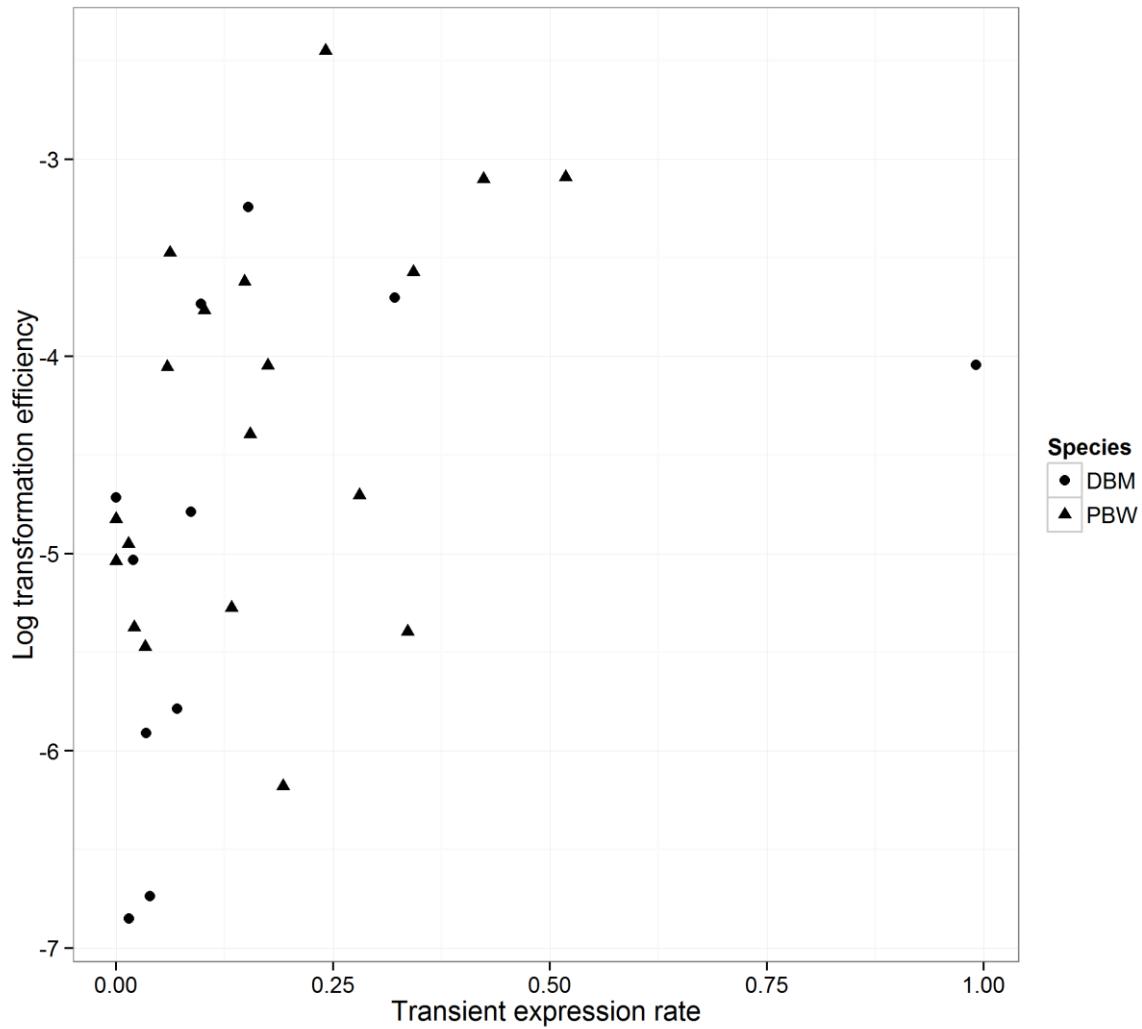


Figure 16. The proportion of injection survivors that display transience (number of transient injection survivors / number of injection survivors) against the log proportion of total injection survivors that give rise to transgenic offspring (number of independent lines / number of injection survivors). PBW - triangles, pink bollworm; DBM - circles, diamondback moth. Transience rate does not appear to correlate well or act as a reliable predictor of transformation efficiency. The distribution appears to be heavily skewed towards zero for the transience rate.

Of the 31 constructs, for which there were data pertaining to the parentage of each unique insertion event, a total of 130 transgenic lines arose (Figure 17).

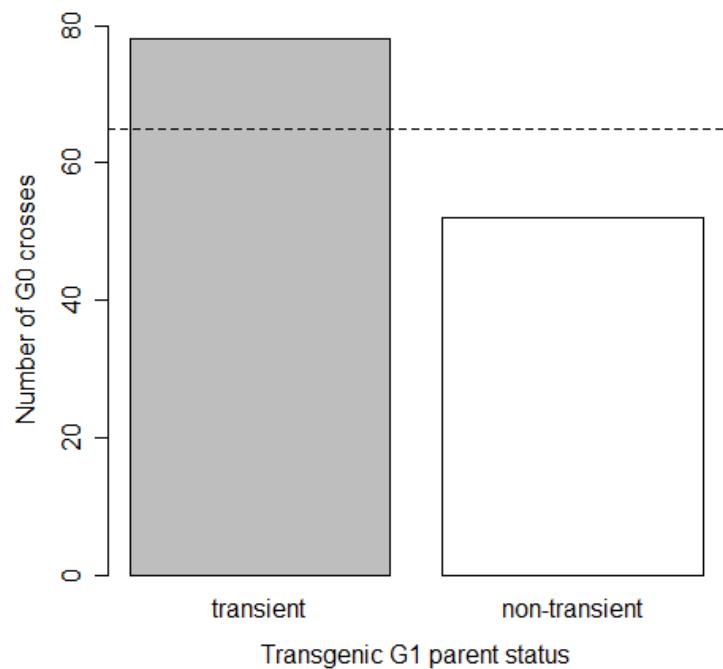


Figure 17. The observed number of transient and non-transient G₀ that sired transgenic G₁ compared to the null model that there would be no difference (1:1). Data were available for 35 unique constructs and 130 unique insertion events. A chi-square test was conducted on the observed (78 transient, 52 non-transient) and tested for a difference to the expected (65 transient, 65 non-transient) indicated by the dashed line. The evidence for the null model is significantly weak ($\chi^2 = 5.2$, $p = 0.0266$), suggesting a real difference. The data is from transgenesis in the Lepidoptera; *Plutella xylostella* and *Pectinophora gossypiella*.

The difficulty in interpreting this data is that the methods applied to crossing the G₀ varied depending on the transience status and, as such, methods were not controlled. As a result, the difference from the null model could be explained by the difference in how they were pooled between the transient and non-transient G₀, rather than the alternative explanation that transgenic offspring are more likely to be produced in transient crosses.

A double-blind test with equal methodology could have been performed to determine if any real difference exists. The difficulty in controlling variables and the large number of replicates to detect a relatively small discrepancy in effect size renders this clarification prohibitive. Instead, the mechanism behind transience was investigated. An experiment was designed to determine whether transience can be caused by construct fluorescent genes being expressed episomally.

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

This was investigated by injecting embryos with two treatments in a double-blind experiment, one group with construct and helper plasmid, the other with just the construct. Any differences in transience rate were examined and the emphasis placed on transience by some researchers would be given an empirical basis. Two negative controls were also included, one with just helper mRNA and the other with just the injection mix and no nucleic acids.

3.3 WHAT IS REQUIRED FOR TRANSIENT EXPRESSION?

Reading several theses completed by students at Oxitec in collaboration with the University of Oxford, it became apparent that there was a belief in an association between presence of a transposase and construct, giving rise to higher transience rates relative to just the construct being microinjected. An alternative explanation is that researchers do not trust their injections and require visual confirmation that the plasmid has entered the embryo. Either way preferable treatment of transient G₀ is defensible. The theses indicate speculation rather than outright belief:

'Inclusion of helper mRNA may result in higher transient-expression ... and therefore higher transformation efficiency.' (Morrison, N., p.137)

Again a direct quote, this time from the literature, reveals this belief may be held by researchers in this field:

'...transient-expression of Tcv in the G₀ parent can be used to preselect G₀s that are more likely to produce transformed progeny.' (Lorenzen, 2002)

Based on this speculative belief and exploration of the injection data an experiment was designed to investigate what ingredients included in an injection mix are associated with transient-expression of the marker, specifically; the construct, the helper plasmid and all combinations thereof.

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

It was assumed that only the construct containing combinations will result in transience (and transformants). Any other treatments that result in transience will be unexpected and therefore interesting. The experiment should also elucidate the effect of the transposase on the transience rate relative to injecting the construct only.

More specifically, the interest lies in whether one of the hypothetical benefits of prioritising transient-expressing pupae for setting up G₁ crosses is that we are seeing expression from somatically transformed patches of cells – we would therefore expect that these patches would occasionally be present in the germline. However, it is possible that transience may be occurring directly from the un-integrated plasmid, thereby reducing the accuracy of transience as a proxy for integration.

3.4 POWER ANALYSIS AND BLINDING

The rigorous blinding was enforced, ensuring that experimenters were ignorant of the treatment being applied when injecting and the treatment that had been applied when screening for transience. A power analysis (Figure 18) was used to determine the number of embryos to be injected to ensure that a success (transient embryo status) would be achieved, given it were possible, with 95% confidence. A simple Poisson model — as an approximation to a binomial model — was used with the probability of success the median transient expression rate from the Oxitec diamondback moth data (compiled by the author from A. Walker's Oxitec lab-books in 2012/2013).

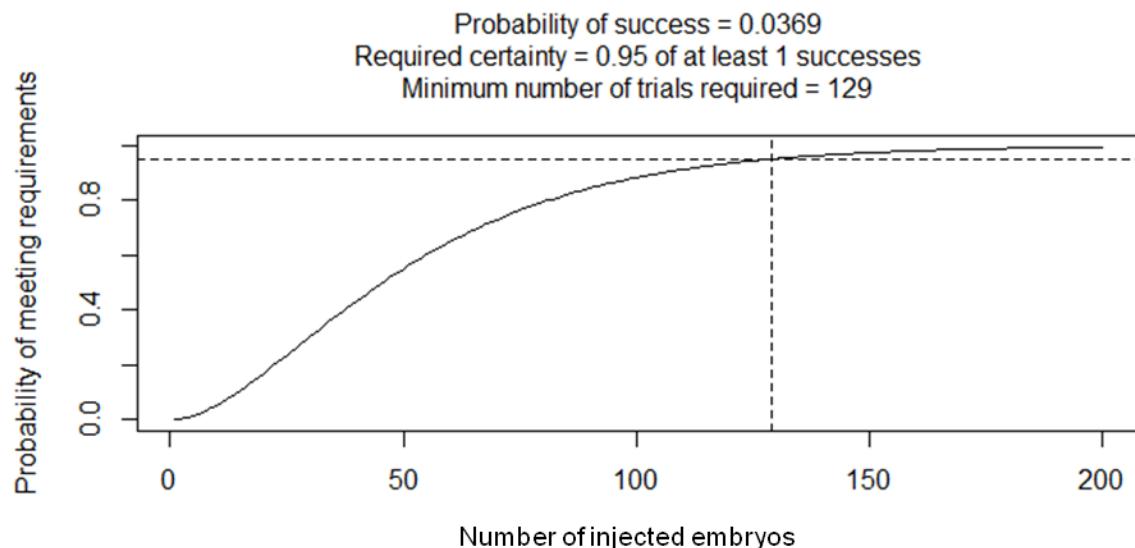


Figure 18. A Poisson distribution defined by the transience rate $\lambda = 0.037$ for a given number of independent trials (G₀ survivors). The number of embryos to be injected and then their progeny screened to be 95% sure of obtaining at least one transient G₀ is 129 in diamondback moth.

Given prior information on transience rates following injection of embryos the model suggests a minimum of 129 injection survivors per treatment. This number is the minimum required to be confident in revealing any qualitative differences between treatments and the transience rate (i.e. getting at least one transient-expressing G₀). Given the median survival rate to G₀ in diamondback moth of 0.23, approximately 500 injections for each treatment are required.

3.5 METHOD

All mixes were made up to a volume of 30 μ l in a micro centrifuge tube. The construct OX4804 (Hr-*ie1*-Amcyan marker) and the helper mRNA OX3081 were used at 500 ng/ μ l and 350 ng/ μ l as standard with 3 μ l of injection buffer (10x) and the rest of the volume made up to 30 μ l using MilliQ water (six aliquots were taken for storage at -20°C). The injections were carried out blind [by Gregory, M. (MG)] as was the screening [Walker, A. (AW) and Harvey-Samuel, T. (THS)] and analysis of results (MG). The treatment assigned to each slide of embryos was randomised by injection mix. Needles and pipettes were changed as appropriate to avoid contamination.

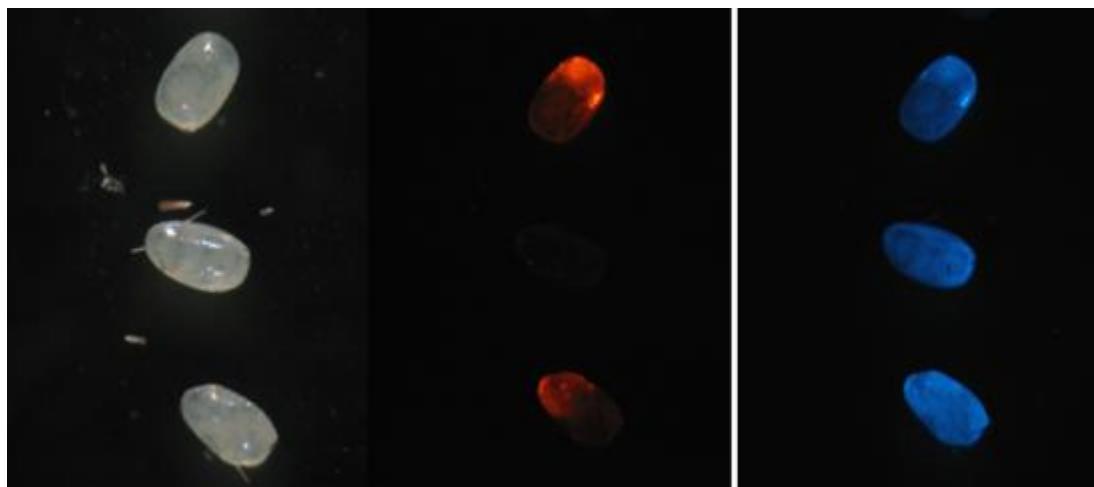


Figure 19. An example of transient expression of a construct marker (OX4673) and helper (DNA, OX265) plasmid injected into diamondback moth embryos. The images are in white light, DsRed and AmCyan filter conditions (from left to right). The middle embryo is not showing transient-expression. The construct contains both DsRed and AMcyan coding regions. Photographs courtesy of Tim Harvey-Samuel.

Embryos were screened for fluorescence (AmCyan) at days one, two and three after injection and scored for transience. Screening of slides with injected embryos was carried out by at least two screeners where possible (to compare screener agreement on subjective nature of transient-expression status). Follow-up screening was carried out by the same experimenter recording results in a dedicated folder kept separate from other experimenters. Larvae were not tracked — by egg transient status — but labelled by treatment received. Following pupation, injection survivors were screened and scored.

3.6 RESULTS

Analyses were carried out blind; labels were used as placeholders until the identity of each treatment were revealed post analysis.

3.6.1 DIFFERENCES BETWEEN TREATMENTS

The probability of transience in an embryo was not constant between treatments. Transience was only detected in embryos that had been injected with either Construct or Construct &

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

Helper treatments. To avoid pseudo-replication data were sub-setted into embryos screened by the same individual to allow comparison (Table 6). A total of 993 embryos were injected.

Table 6. A stacked two-way table showing the number of eggs screened by two experimenters, the treatment injected and the following transience status of the embryos. The embryos screened were not necessarily identical due to differences in attendance between colleagues. Screeners did not screen exactly the same number of slides and embryos thereon. The embryos injected with helper only were only screened by one experimenter. All screening was carried out blind to knowledge of the treatment. Screeners were aware of previous transience status they had assigned to an embryo due to the recording method. Screener 1 – AW, Screener 2 – THS.

| Screener 1 | | | | |
|-------------------|------------------|--------------|---------------|-------------------------------|
| | Construct | Water | Helper | Construct & Helper |
| Non-transient | 144 | 352 | 88 | 200 |
| Transient | 94 | 0 | 0 | 135 |
| Screener 2 | | | | |
| | Construct | Water | Helper | Construct & Helper |
| Non-transient | 134 | 370 | 0 | 117 |
| Transient | 97 | 0 | 0 | 111 |

Embryos were screened at least once for transience for the first 3 days post-injection. The transient status of each embryo was assigned as either transient (transience detected on at least one of the three days) or non-transient (no-transience detected on any screenings). The probability of transience was inspected for each experimental unit. After statistical validation of pooling (Xu *et al.*, 2010), data were pooled by treatment and summarised (Figure 20).

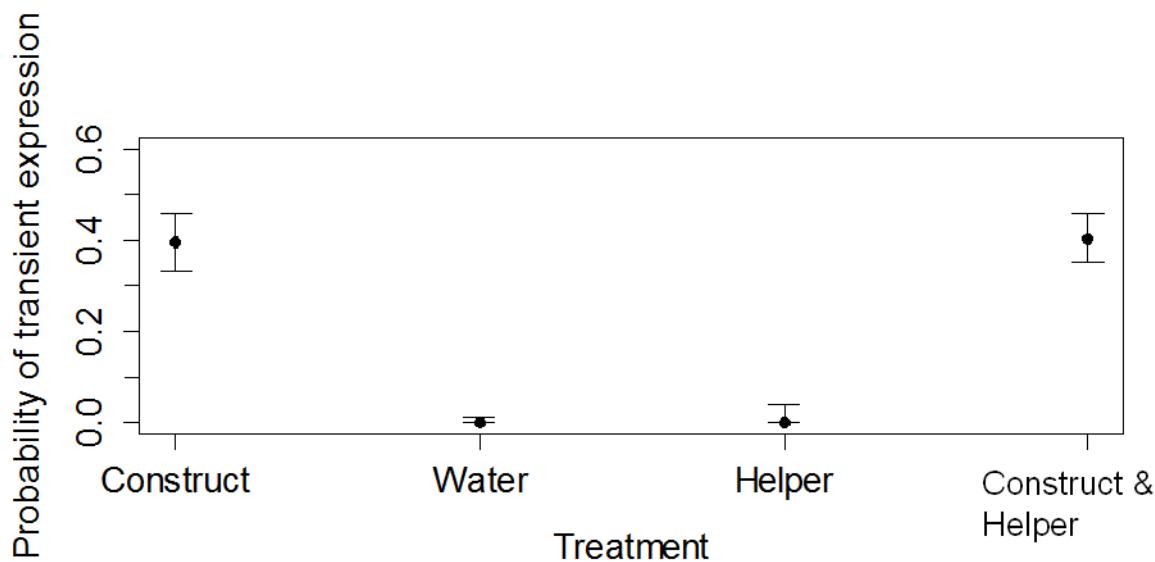


Figure 20. The probability of an embryo testing positive for transience screening 1-3 days post injection given different injection mixes. The mean was calculated after pooling individual embryo status data from each slide – the total number of embryos showing transient-expression divided by the total number of embryos screened per treatment. Binomial confidence intervals using the Klopper-Pearson method at the 95% level were also added for each treatment (package "binom" in R by Doraj-Raj, 2009). An injection mix with no added nucleic acids (water) and an injection mix of just mRNA helper produced no embryos identified as positive for transience. Both treatments containing the construct had similar transient-expression rates of about 40%.

Injecting embryos with treatments that did not contain the DNA construct did not produce any transient embryos. Comparing water and just helper suggests that contamination of the injection mix with constructs containing AmCyan did not occur and may be unlikely under normal experimental conditions. Furthermore, it also suggests that auto-fluorescence that develops in dead embryos is unlikely to be mistaken for transience by experienced screeners.

The remaining treatments did produce transience; the construct was necessary and sufficient for a positive transience status in injected embryos. The construct and helper combination had a slightly higher mean transience probability compared to just the construct treatment, with confidence intervals showing considerable overlap.

The transient expression statuses of embryos were modelled using a generalised linear mixed effect model (GLMM) with binomial error distribution (Bates, 2012). The maximal model was fitted and then model simplification proceeded in a stepwise manner towards the minimum

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

adequate model, elucidating the important variables in explaining the variation in the data (additional detail provided in Appendix 9.8).

The fitted means of the model were extracted (Table 7). The parameter values represent differences between means, so to get the transience rates, parameter values were added to the intercept before back transforming (Crawley, 2005). The fitted model returned:

Table 7. Parameter values from the generalised linear mixed effect model: transient status ~ treatment + (|slide), family = binomial where tilde zero means equivalent to zero. Comparison between data and model reveals an adequate fit.

| | Construct | Water | Helper | Construct & Helper |
|--------------|------------------|--------------|---------------|-------------------------------|
| Model | 0.399 | ~0 | ~0 | 0.398 |
| Data | 0.395 | 0 | 0 | 0.403 |

The z-test statistic suggested that the treatments that included the construct were not significantly different from one another. Similarly the treatments without the construct were not dissimilar. A new factor was created from the original 4-level factor; it was condensed into a new fixed effect factor with 2-levels by combining results into the binary variable construct status (“Construct present” or “No construct”). The random effect of slide was still included (lmer (transient status ~ construct_status + (|slide), family = binomial)).

The model was compared with the 4-level treatment model using analysis of deviance. There was no significant loss in explanatory power so the model was adopted ($X^2 \sim 0$, df=2, p~1). The new model, which was now the minimum adequate model, provided a transformed parameter value for the probability of transience in an embryo for injection mixes with the construct as 0.398. The embryo transience rate of 0.40 compares to Mohammed & Coates (2004) rate of 0.57 in the potato tuber moth, *Phthorimaea operculella* Zeller (Lepidoptera: Gelechiidae).

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

The non-negligible effect of the slide on the model may be due to the limited number of experimental units and/or environmental variation during injecting, as needles were loaded with a randomised treatment often consecutive slides would be injected with the same treatment to save on waste. This aspect of the trial was not perfectly randomised, as each slide did not randomly receive a treatment, as it was dependent on prior slide treatment. This analysis shows why it is suspect to ignore the random effect, in this case there is a correlation between the data coming from each experimental unit or slide (albeit a small effect), questioning the applicability of the inferred sample statistics to the population. There could be a relationship between receiving a transience-inducing dose and the survival of the embryo; perhaps those embryos that are pierced but not injected with an appropriate volume are more likely to survive, for example.

At the resolution of the transient status of the embryo, the screeners did not vary in their assessment of embryos elucidated by the modelling process described above. However, this says nothing about the day-to-day screening results of specific embryos on one (t_1), two (t_2) and three (t_3) days after injection, as individual interpretation of transient statuses were recorded but will be discussed later.

When the response variable can only take on a limited range of values the utility of diagnostic plots is decreased. The nature of the mixed effect model makes assessment and biological interpretation difficult. However, the main utility of the model was to suggest that the addition of helper into the injection mix does not significantly change the transience rate in embryos. Therefore refuting the inference that transience requires both construct DNA and helper mRNA to have successfully made it into the embryo. Further, the addition of helper shows no interaction (no synergy) with the construct in improving the transience rate in those embryos studied.

3.6.2 DIFFERENCES IN EMBRYO TRANSIENCE STATUS BETWEEN CONSECUTIVE DAYS

Transience status — defined above — loses some information recorded during the experiment by condensing down the day-to-day embryo status to an overall status of whether the embryo was ever transient on any of the 3 days post-injection. Is this likely to bias results? How does the transient status vary from day to day, is it dynamic or constant? Does the rate of change vary between days or is it constant? Is it different between treatments?

Here we shall formulate a simple stochastic model for the transient status of the embryo on a particular day after microinjection of the construct (and helper) using a simple Markov Chain model. Each time step is considered discrete and represents approximately 24 h since the status was last assessed for the specific embryo by THS. The reason this analysis was chosen was each embryo state has serial dependence through time. This means that the current embryo transience state only depends on the previous state, and not on earlier states. This is a simplifying assumption that is often made.

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

Our model contains two states: transient or non-transient. At each time step the allowed transitions are (where \rightarrow can be read as “then”):

- transient \rightarrow transient
- transient \rightarrow non-transient
- non-transient \rightarrow non-transient
- non-transient \rightarrow transient

For simplicity and mathematical tractability the binary state of the embryo is described by the random variable X_n and assumed to be independent and identically distributed.

Where $X_n = (1 \text{ or } 0, \text{ transient or non-transient})$.

The probability of transition between states is given by α , β , γ and δ (Figure 21).

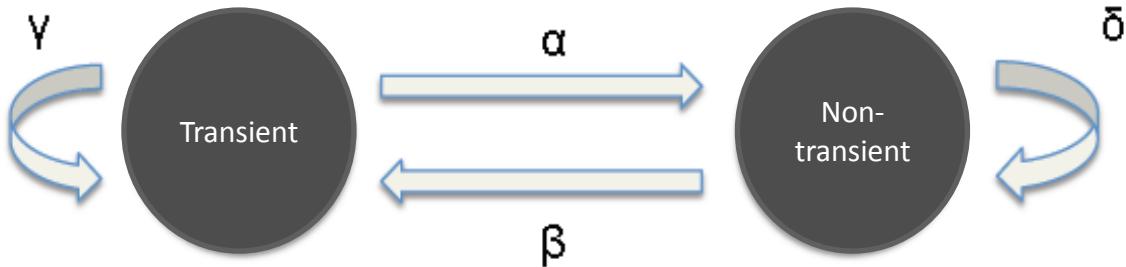


Figure 21. The Markov Chain model describes the transition between states — transient and non-transient — with the possibility of staying in the same state through discrete time periods t_0 , t_1 , t_2 and t_3 . The probability of transition between states is given by α , β , γ and δ .

Therefore between time-steps t_0 and t_1 only transitions from the non-transient state are possible.

For the microinjected embryo-treatment of construct-only (OX4804) with screener THS, the initial states and transitions are given by a Markov Chain model (Figure 22).

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

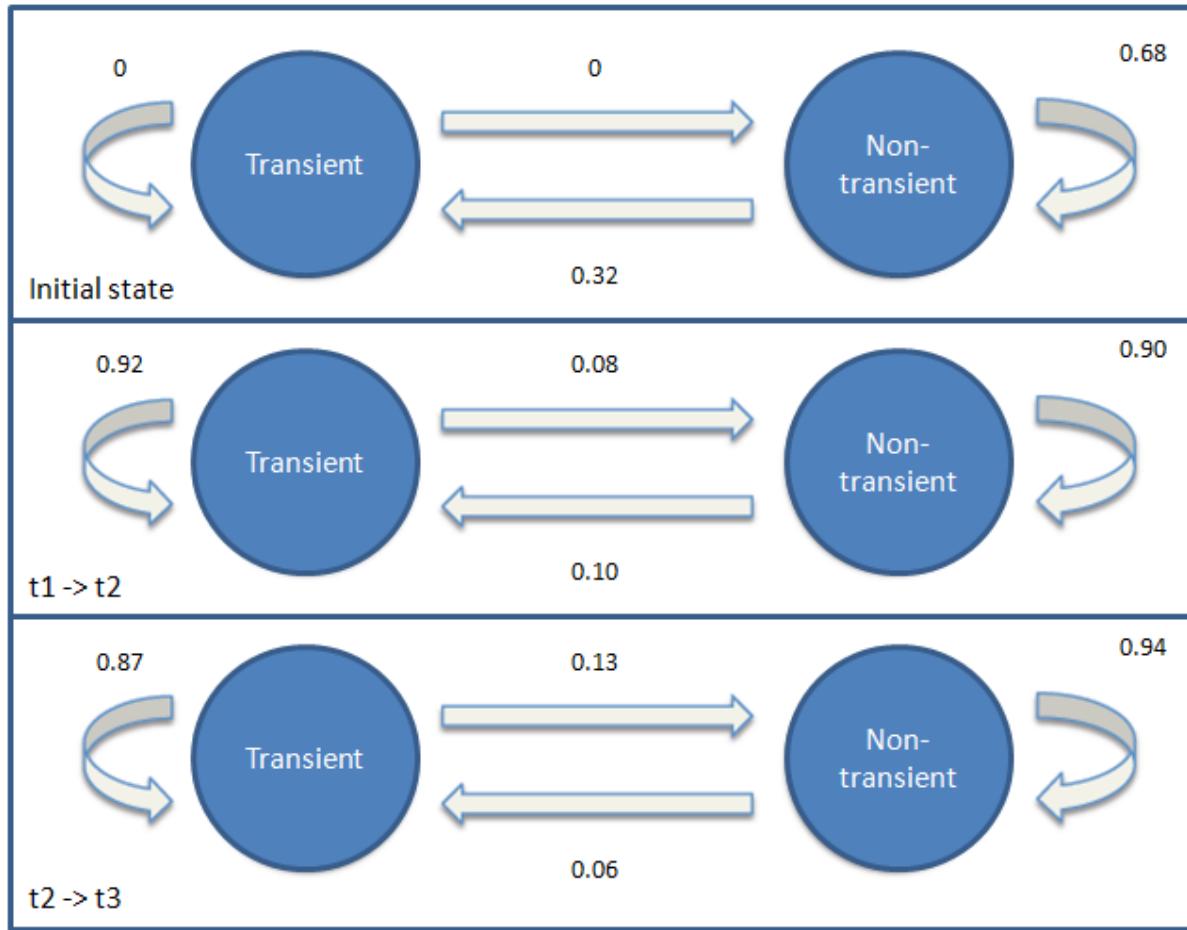


Figure 22. A simple Markov Chain model was used to describe the state of embryos and the probability of state transitions between days for the screener THS. The initial transition — t_0 then t_1 — is a special case as all embryos start as non-transient. Arrows represent transitions with the proportion of embryos transitioning between states given by the adjacent probability. The state of an embryo between t_1 and t_2 is modelled by $t_1 \rightarrow t_2$, the state of an embryo between t_2 and t_3 is modelled by $t_2 \rightarrow t_3$.

The initial state tells us that only about 30% are likely to be transient and 10% of these per day will be assessed to be non-transient thereafter. The probability of transition between states appears approximately constant, with a maximum discrepancy of 0.05 between the transition matrices at $t_1 \rightarrow t_2$ and $t_2 \rightarrow t_3$. This interpretation ignores the complication that initially there will be more non-transient embryos due to the uneven probability of each state initially. Therefore, one day after injection — t_1 — may not be the optimal time to assess transient status if it is to be used as an indicator by researchers.

For the microinjected embryo treatment of construct and helper (OX4804 and OX3081) with screener AW the initial states and transitions are given by the model below (Figure 23).

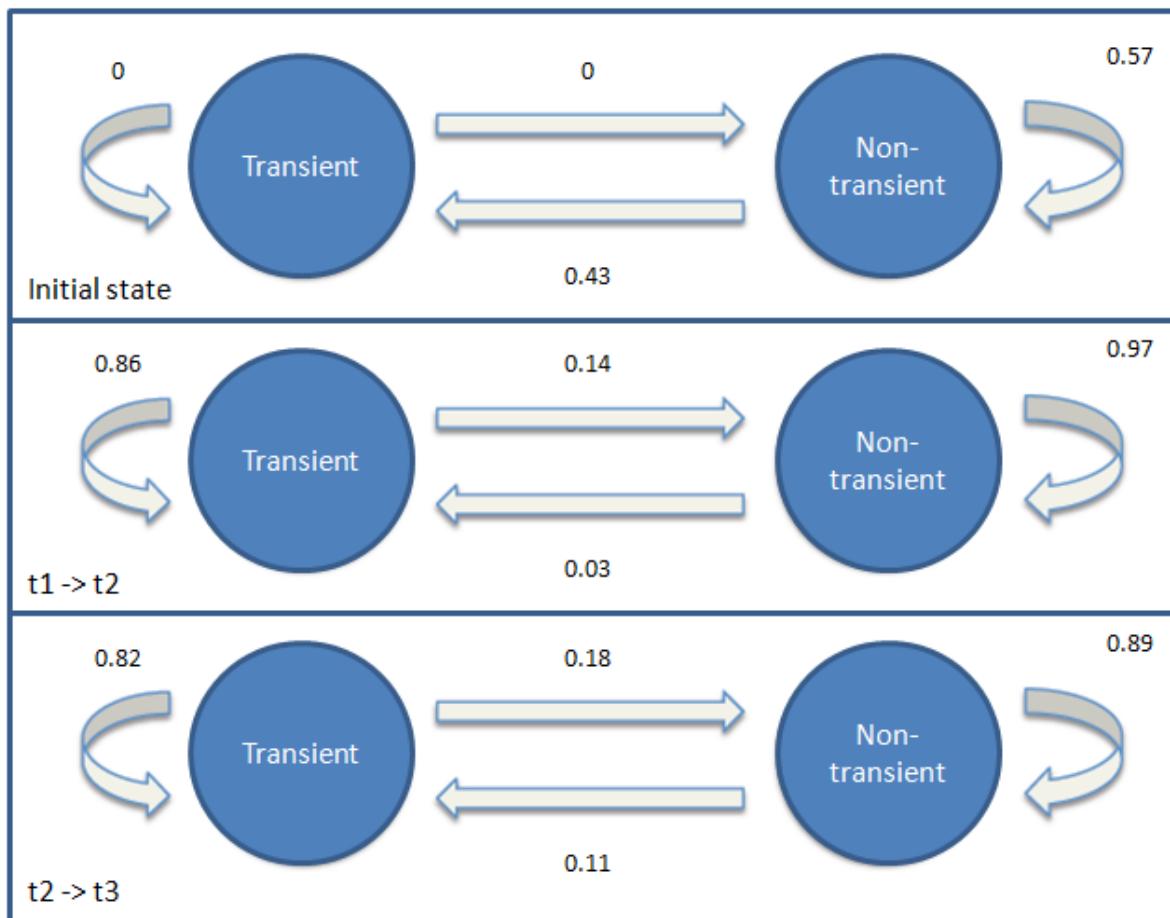


Figure 23. A simple Markov chain model was used to describe the state of embryos and the probability of state transitions between days for the screener AW. The initial transition — t_0 then t_1 — is a special case as all embryos start as non-transient. Arrows represent transitions with the proportion of embryos transitioning between states given by the adjacent probability. The state of an embryo between t_1 and t_2 is modelled by $t_1 \rightarrow t_2$, the state of an embryo between t_2 and t_3 is modelled by $t_2 \rightarrow t_3$.

The transient state showed similar stability between days to that observed by THS. On day two the probability that a randomly chosen embryo was transient was 39%, a slight decrease. Of the 91 transient embryos out of the total 228 microinjected, 71 displayed transience through time periods t_2 and t_3 . With this treatment/screener combination day one was the time with the highest probability of assessing a randomly chosen embryo as being in the transient state.

3.6.3 WHEN SHOULD I CHECK FOR EMBRYO TRANSIENCE?

Using these parameters derived from the data, a Markov Chain Monte Carlo simulation was run to ascertain the 'best' day to check the transient status of an embryo (the day that most

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

consistently had the highest number of transients). The stochastic nature of the model results in different outcomes given identical inputs, however it was observed that t_2 always had fewer transient embryos than t_1 or t_3 .

Accordingly, the difference between the numbers of transient embryos on t_3 minus t_1 was used as an indicator variable describing the direction and magnitude of the change. If there were more embryos that were transient on day one compared to day three that produced a negative number for the difference (if $t_3 - t_1 = -R$, then $t_1 > t_3$). A positive difference was given if the opposite were true and there were more transient embryos on t_3 compared to t_1 ($t_3 - t_1 = +R$, then $t_1 < t_3$). The simulation was run 10,000 times and a histogram of the results plotted to determine the average behaviour of the model (Figure 24).

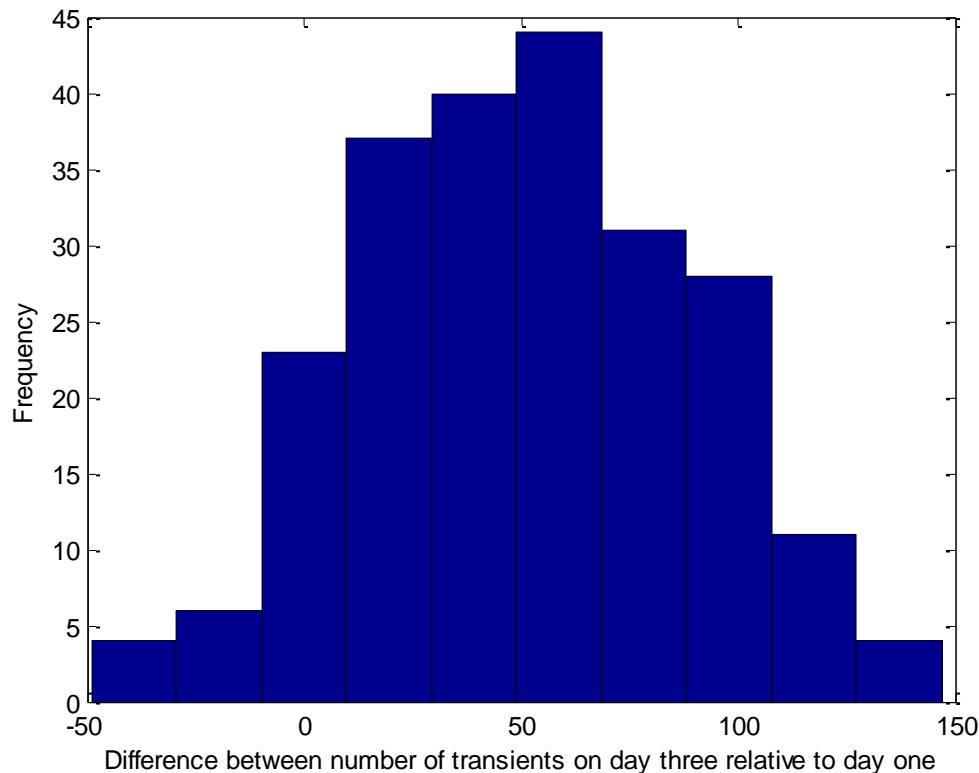


Figure 24. A histogram of an indicator variable (the sum of the number of transient embryos on day three compared to the sum of embryos on day one) frequency distribution over 10,000 simulations of a Markov Chain model of the experimental process.

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

The interpretation of the positive skew and location of the mean, median and mode ($>> 0$) suggests that a greater number of simulations had t_3 as the time step with most transient embryos relative to the initial number on t_1 . This suggests that researchers should assess for transience 3 days after injection to increase their chances of seeing the most embryos positive for transient expression. This knowledge is useful when doing a transposition assay in a hitherto untransformed insect using *piggyBac* and a fluorescent protein-promoter combination that is untested (Mohammed & Coates, 2004). This interpretation is only valid for the specific situation described above given the parameters of the transition model derived from the OX4804 injection data. It seems plausible that the parameter estimates may be true for constructs that use the Hr5-*ie1*-AmCyan transformation marker in diamondback moth but may not apply to other promoter and/or fluorescent protein combinations.

The conclusion may be confounded by the effect of death on embryo status with some reports of auto-fluorescence occurring quickly after death. Furthermore an embryo that may have been non-transient on day one may have developed transient status in the proceeding days had it not died. This could be avoided by using a multiple state Markov Chain model with appropriate coding for the embryo status including dead or alive. An embryo could only be alive and transient during this experiment.

3.6.4 PUPAL TRANSIENCE

Transience statuses at the pupal stage were compared between treatments and were given as the number of successes — transient-status positive pupae — versus the total number of pupae screened per treatment (Table 8).

Table 8. The total number of transient pupae given a total number screened determined by the number of insects surviving injection and pupating by treatment. The construct was necessary and sufficient for transient-expression in pupae. Transience was detected at a lower rate in pupae compared to injected embryos. Data shown in parentheses are from a secondary experiment that was run for the helper-only treatment to remedy the accidental destruction of the first experiment's pupae treated with helper only described in the main body of the text.

| Treatment | Transient pupae | Total pupae screened |
|----------------------|-----------------|----------------------|
| Construct only | 1 | 50 |
| Water | 0 | 112 |
| Helper only | 0 (0) | 1 (88*) |
| Construct and Helper | 2 | 53 |

The pupae transience status data reinforces the idea that the construct is necessary and sufficient to produce transience in pupae. The probability of a pupa being classed as transient is much reduced relative to that of the embryo. This suggests differential mortality or the "transient" nature of transience as well as the difficulty of detection fluorescence through cuticle may reduce transient-expression detection rates.

It could not be ruled out that transient pupae did not result from injection of either of the controls due to insufficient power, as outlined in the power analysis. This was due to insufficient injections as well as complications with rearing resulting in high larval mortality. However, at a later date 217 injections of 'Helper only' were conducted giving rise to 88 G₀. No transience was detected. This improvement in power compromised the blinding of the experiment.

The pupal transience rate of the construct and helper injection treatment, 0.038 (2/53), is similar to previous estimates of 0.037 and estimate of 0.021 (15/683) (Harvey-Samuel, 2014). This was an order of magnitude lower than the embryo transience rate.

3.7 CONCLUSIONS

The experiment showed that for transience to occur only the construct is necessary and sufficient when AmCyan is driven by the Hr5-*ie1* promoter. This suggests that if a researcher wants to test a given promoter-flourescent-marker combination one can inject just the construct for a transposition assay screening at day three post-injection. Insertion of the construct with the assistance of the helper transposase is not necessary for transient-expression. This is unsurprising as previous work using transposition assays had shown immediate expression upon injection when using Hr5-*ie1* (Mohammed & Coates, 2004). Interestingly, the helper does not appear to increase frequency of transient expression in embryos.

Transience occurs — or is detectable — an order of magnitude less frequently in pupae compared to embryos (up to 3 days post-injection). This suggests that many embryos which show transient expression do not necessarily show it at the pupal stage: that transience is transient even within the lifetime of the insect. This may be related to the restructuring of the insect during metamorphosis. This suggests those researchers that use pupal transience as an indicator of the construct, as a minimum, being successfully injected are overlooking many embryos which showed transience but no longer do as pupae. Furthermore, the data suggest that transient-expression does not imply insertion, as transience was detected without the helper transposase.

Tracking individual embryos revealed that transience is a dynamic state, with the day of assessment affecting the estimate of transient embryos by up to 10%. This was quantified and described by a Markov Chain model. The 10-fold decrease in transience from larva to pupae may be explained by the increased difficulty in detecting the signal from a larger and more opaque life-stage. This could be assessed by comparing phenotypic assessment of transience with a genetic test by using PCR to assess whether the insect is positive for the plasmid.

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

The experiment can be used to inform decision-making regarding whether to discard non-transient embryos and just use transient insects for setting up crosses. The results suggest that transient expression indicates functional plasmid is present rather than somatic transformation of the embryo, however we cannot infer whether transient expression is meaningful to the injector. A large-scale controlled experiments with a huge number of injections would be required to detect a difference in effect size — say 10% — between the transformation efficiency of G₀ with a positive or negative transience status at the embryo stage. Due to the prohibitive scale of this experiment this debate is unlikely to be settled any time soon with most inferences speculative. Rather than tackling the problem in one single massive experiment, the problem could be addressed by incorporating it into routine research, whereby the results of lots of low-powered studies with identical design could be combined into a single summary statistic with a measure of uncertainty, in order to detect a real difference (McShane & Bockenholt, 2013). With coordination between researchers this problem could be resolved quickly.

This chapter has highlighted — as a cutting edge discipline — transforming insects with *piggyBac* has some gaps in knowledge. To correct this, we should incorporate controlled experimental design into our routine research to elucidate some of the areas for which empirical data is lacking and the power of the anecdote reigns. Ultimately it reminds us of the caveats to bear in mind when drawing any conclusions about descriptive or observational data (Petitti, 2004):

1. Do not turn a blind eye to contradiction.
2. Do not be seduced by mechanism.
3. Suspend belief.
4. Maintain scepticism.

These points, combined with unestablished validity of the data, remind us to be careful in our conclusions. This does not preclude any inferences or increase in knowledge of insect transgenesis. For example, although observing A then B does not mean A causes B. Nevertheless,

Chapter 3 | What components of a microinjection mix are required for transient expression of a marker?

if A causes B, then an association should be expected in the data. Hence, controlled experiments are the only means to elucidate a causal link.

CHAPTER 4: ENGINEERING MALE STERILITY IN DIAMONDBACK MOTH (*PLUTELLA XYLOSTELLA*)



4.1. INTRODUCTION

4.1.1 AN OVERVIEW OF THE PATERNAL EFFECT SYSTEM

Lepidoptera are significant pests of food and fibre crops with the emergence and persistence of insecticide resistance compromising conventional treatment strategies (Simmons *et al.*, 2010). SIT in the Lepidoptera provides an additional weapon in the arsenal of integrated pest

management (Marec *et al.*, 2005). In SIT, the target wild females must, somewhat paradoxically, mate with the mass-reared sterile males to achieve reproductive failure, owing to the non-viability of eggs produced by such crosses (Perez-Staples *et al.*, 2013).

Radiosterilisation is still considered a viable method in some species (Vreysen *et al.*, 2010), with improvements in sterilisation technology efficacy and safety (Helinski *et al.*, 2009) rekindling interest and research effort. This technique, however, has been historically difficult when applied to lepidopteran species as they have proven resistant to radiation damage, possibly due to the structure of their chromosomes (Marec *et al.*, 1999). The diffuse arrangement of lepidopteran centromeres may offer resistance against imposed genetic sterility based on double-stranded breaks of DNA. Ultimately, radiosterilisation is limited in its species applicability, and does not provide a consistent method of producing competitive sterile males in the Lepidoptera.

This project aims to replace atomic radiosterilisation with a genetic means of introducing double-stranded breaks into the germ line DNA of the reared males, thereby negating the negative side-effects on the sterile male associated with radiosterilisation (Alphey *et al.*, 2006). Sterility will be achieved by inserting a genetic construct into the germ-line using the *piggyBac* transposable element (Handler, 2000). This construct will incorporate components that contribute to a repressible control of sperm viability; including the transgenesis marker module and the sperm-specific module (Figure 25).

The development of such a sperm-lethal system should be reliant on the expression of an endonuclease during a specific window of sperm development to create the double-stranded breaks in an otherwise superficially normal sperm phenotype (a transgenic sperm capable of fertilisation of eggs in the female). This chromosomal shredding should result in non-viable offspring that die at the early embryo stage. To facilitate transport to the DNA and binding the endonuclease should be fused to a chaperone molecule.

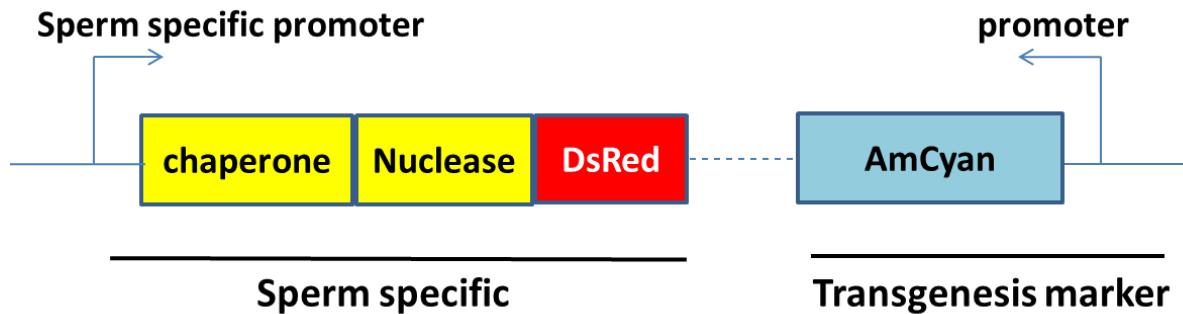


Figure 25. The “Paternal effect” system consists of a generic transgenesis indicator module and the sperm-lethal module. The generic promoter and AmCyan will allow screening of G₁ for transgenesis. The paternal effect module will express a DNA cutting nuclease and DNA binding chaperone causing male sterility given expression at the appropriate time under control of the sperm-specific promoter. For clarity, the components for repressibility of the RIDL binary system are omitted from this simplified schematic. The nuclease and DsRed2 will also not be fused but can be separated at translation using ubiquitin.

The system should be repressible so that insects can be mass reared and used for an SIT programme as required. This could be achieved using the dietary control of the *tetO-tTAV* binary system (Hara *et al.*, 2009). The paternal effect system also avoids current controversy over the hazard of environmental tetracycline jeopardising the reliability of conventional RIDL systems.

The system requires detailed knowledge of the process of spermatogenesis in the target insect species. Although this may be lacking in our two species of interest, in the model species, *Drosophila melanogaster*, it is well-characterised and shares many features.

Sperm-lethal in this thesis refers to sperm carrying dominant lethal mutations that lead to death of the sperm and egg combination. This could be due to strict paternal effect lethality, such as a failure of the acrosome reaction, or other phenotypes affecting fertilisation, or due to early-acting zygotic lethality caused by a factor principally active in the sperm, as seen for example in irradiated sperm. The sperm-lethal and paternal effect systems encompass all of these.

4.1.2 SPERM DEVELOPMENT IN THE LEPIDOPTERA

Insects vary greatly in their reproductive anatomy, behaviour and physiology but there are shared characteristics (Chapman, 2012). The male genital system in insects consists of two

testes. At the most basic level the testes are an organ specialised in cell differentiation from germline stem cell to mature sperm capable of fertilisation.

Each testis is generally composed of several testis tubes or follicles, each acting as a discrete unit of sperm production, which open into a shared deferent duct, the vas deferens (Figure 31). In the Lepidoptera these follicles can be incompletely separated from one another by a thin follicle epithelium. Along the vas deferens - often a large compartment - the seminal vesicle occurs, that in turn opens into the ejaculatory duct. Near the distal end of the ejaculatory duct a system of accessory glands is present. The copulatory organ, the size and complexity of which can vary greatly between species, is located at the terminus of the genital system (Dallai, 2014).

At the apex of each follicle are germ cells located in a proliferative centre. Germ cells are progressively displaced and move further from the germarium or proliferative "hub" complex to the distal end of the follicle; undergoing mitotic and meiotic divisions as they progress.

Thus the entire range of developmental stages of sperm cells can be observed in the follicle, with the earliest stages found distally in the germarium and the most sperm-like in the proximal part of the follicle adjacent to the vas deferens with a developmental gradient assumed in-between. This progression has traditionally been designated into different stages defined and characterised by cell types and ploidy. We can combine knowledge of *Drosophila melanogaster* spermatogenesis with diamondback moth follicle structure to create an idealised hybrid-diagram (Figure 26).

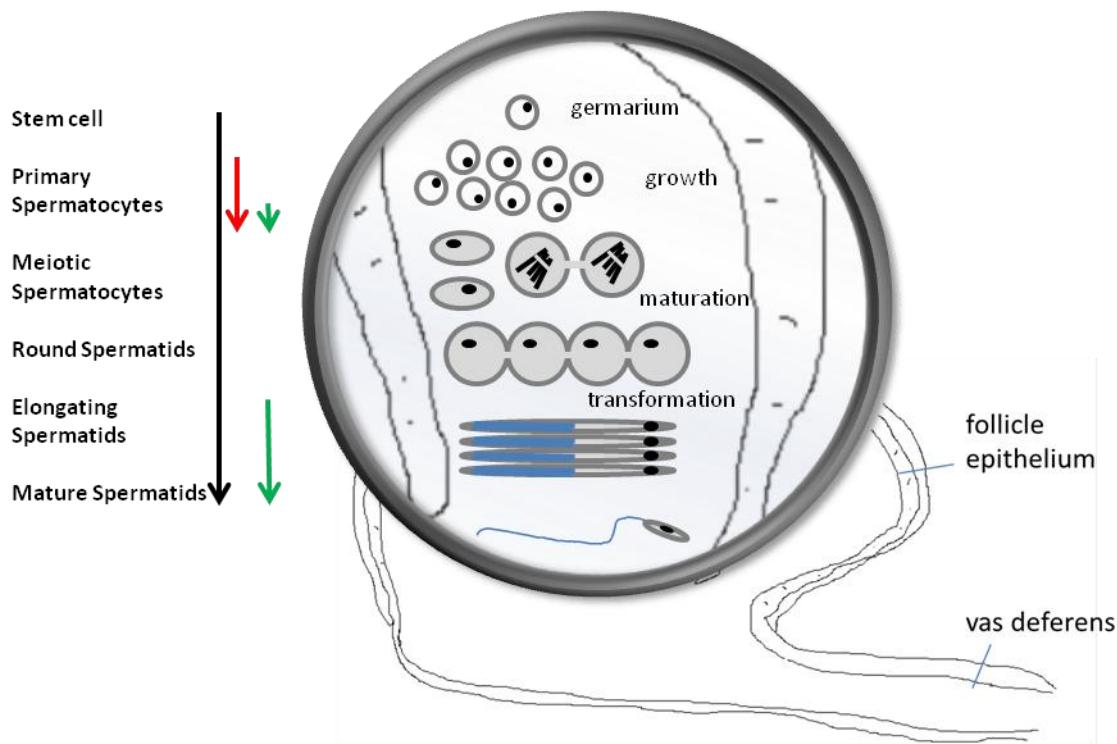


Figure 26. Diagram of the different stages of spermatogenesis in *Drosophila melanogaster* framed by the anatomy of a *Plutella xylostella* (After Dorogova *et al.*, 2009; Jin, 2011; Chapman, 2012) testis. The different stages of growth can be grouped into the germarium-based divisions, growth and massive transcription associated with spermatocytes (the largest sperm cells), genetic maturation and haploidy and finally transformation into the distinctive sperm phenotype. The development tends to be modular with cells packaged in cysts. Notice the development gradient through the follicle with age and maturation stage of the cell corresponding to its position. The timing of transcription (red arrow) and translation (green arrow) of *Ceratitis capitata* $\beta 2$ -tubulin is shown (Jin, 2011).

The process of spermatogenesis involves initial growth and mitotic cellular division, meiotic division to produce the spermatids and finally a transformation from spermatid to spermatozoa. These different transitional stages initiate at different developmental life-stages of the organism.

Spermatogonia are immature germ cells that undergo a series of mitoses to give rise to a large pool of cells that become committed to entering meiosis. The purpose of meiosis is to halve the genetic material and promote the crossing-over of genetic information between homologous parental chromosomes generating genetic diversity in spermatozoa. Following meiosis the cell must continue to differentiate from a haploid tugboat-like-spermatid to a motile streamlined-torpedo-like-spermatozoan.

The mature sperm are released from the testis into the upper vas deferens as early as the pharate adult in some Lepidoptera, albeit with a circadian rhythm (Kotwica *et al.*, 2011). In some species the sperm are ready for transmission to the female, as part of the spermatophore, through the aedagus. However, in some Lepidoptera, eupyrene sperm do not attain full maturation and capacitation until after emigration from the spermatophore into the spermatheca (Swallow & Wilkinson, 2002).

Pivotal to the whole process of spermiogenesis are the microtubules (reviewed in O'Donnell & O'Brian, 2014). Dynamic microtubules are essential for the assembly of various microtubule-based structures that participate in spermatid remodelling, alluded to above. Microtubules are hollow tubes approximately 25 nm in diameter. They are comprised of α and β tubulin heterodimers.

The tubulins show a diversity to match the range of functions performed by these microtubules. Marking tubulins with fluorescent proteins can reveal the role they play in the cell during spermatogenesis. This elucidates the intimate relationship between microtubules and chromosomal reshuffling that occurs during this impressive cellular transformation from diploid germ stem cell to haploid gamete capable of existence outside of the body of the insect and potential fertilisation.

4.1.3 SPERM-SPECIFIC REGULATORY SEQUENCES

Transgenesis to provide fluorescent sperm marking has been achieved in some important agricultural pest species, such as the dipteran Mediterranean fruit fly (Medfly, *Ceratitis capitata*) (Scolari *et al.*, 2008). It has also been proposed that sperm specific regulatory sequences could be used to provide a strategy for genetic sexing or a mechanism for the spread of transgenes through the germline (Smith *et al.*, 2007). In contrast this thesis aims to provide the means to control insect pest populations by introducing double-stranded breaks into the DNA of sperm

cells of released males. This is to be achieved by hijacking transcriptional control of an endogenous sperm specific gene.

4.1.4 THE IDEAL GENE FOR THE JOB

Spermatogenesis demonstrates a dramatic change in cell morphology and behaviour mediated by transcriptional activation and deactivation (Friedlander, 1997; Lehrer-Goldshtein *et al.*, 2010; Kanippayoor *et al.*, 2013). The associated spatial and temporal gene regulation mechanisms could be exploited as a means to regulate an inserted artificial genetic construct in the same manner.

Given prudent promoter choice, the effector gene (or genes) would be expressed at an appropriate time, so that the sperm “appears” normal and functional up and until the point of providing viable genetic material for the zygote. Several cellular structures and their associated genes are transcribed or expressed selectively during spermatogenesis and meiosis in the male gonads of insects (Li *et al.*, 1998; Catteruccia *et al.*, 2005; White-Cooper & Bausek, 2010) offering several options.

In very general terms, genes expressed in testis can be ubiquitously expressed (testis enriched expression), or can have testis-specific expression. The testis-specific category may be preferred for our candidate gene as it would avoid off target effects outside the sperm (in the soma). The candidate promoter should also maintain specificity of expression independent of chromosomal location (Hoyle *et al.*, 1995).

4.1.4.1 B2 TUBULIN

Microtubules are dynamically assembled from α and β heterodimers. This heterogeneity can respond to the diversity of functions in different tissues under various conditions (Kawasaki *et al.*, 2003). This diversity of function mostly arises from the divergent C-terminal sequence (the body of the molecule is also important (Popodi *et al.*, 2008)) permitting the classification into

isotypes, and the ability to undergo numerous post translational modifications (Saussede-aim & Dumontet, 2009). One such functional role is to provide sperm with the motility required for fertilisation.

Motile axonemes are ancient organelles with a long and conserved evolutionary history, which provide structure to flagellae, such as sperm tails. Consistent with this, both α and β forms of tubulin that make up the motile axoneme are conserved in eukaryotes. Tubulins are encoded in multi gene families, each member of which is expressed in a specific tissue and temporal pattern (Kawasaki *et al.*, 2003).

$\beta 2$ -tubulin is distinguishable from other tubulins in the gene family by the motile axoneme-specific sequence motif (Dutcher, 2001). These last 15 amino acids in the protein, which have been identified as isotype-defining, produce a third of the differences in amino acid sequence between $\beta 1$ and $\beta 2$ -tubulin (Nielsen *et al.*, 2001). $\beta 2$ -tubulin is used for the motile sperm tail axoneme and has barely changed in insects in 60 million years (White-Cooper & Bausek, 2010; Nielsen *et al.*, 2011). The axoneme motif most likely mediates central pair assembly through isotype-specific interactions with other proteins (Popodi *et al.*, 2005) and may be tightly constrained by evolution (Nielsen *et al.*, 2006).

In the *Drosophila* post-mitotic male germ cell, $\beta 2$ -tubulin is ubiquitous in non-motile sperm-specific microtubule assemblies, playing a functional role in: meiosis (Davis & Miller, 1988); sperm-head formation (Kawasaki *et al.*, 2003); axoneme formation (Kemphues *et al.*, 1982); and spermatid alignment (Fackenthal *et al.*, 1993). In the post-mitotic germ cells $\beta 2$ -tubulin is the only functional isoform found in *Drosophila* sperm (Kemphues *et al.*, 1982).

The multi-functional nature of the $\beta 2$ -tubulin isoform suggests a high rate of transcription compared to other spermatogenesis related genes (confirmed by transcriptome analysis of RNA-seq, diamondback moth $\beta 2$ -tubulin transcription was testes-specific and had the fourth highest transcript count in diamondback moth testes; personal communication, Gong, 2013).

Interestingly in *Drosophila*, where sperm tails are commonly very long, $\beta 2$ -tubulin comprises 35% of the mass of the axoneme.

$\beta 2$ -tubulin expression (Raff *et al.*, 1982; Davis & Miller, 1988) is likely to occur after all mitotic divisions have been completed (Kemphues *et al.*, 1982) providing a suitable promoter for transgene expression with the desired precision for a temporal-spatial window. The gene was shown to be first expressed at spermatogenesis during late larval development and continually throughout male adulthood in *Drosophila melanogaster* (Zimowska *et al.*, 2009).

4.1.4.2 B2-TUBULIN IN OTHER TAXA COMPARED TO LEPIDOPTERA

Amino acid sequences of the proximal region of the axoneme motif (the -COOH terminus sequence) of the $\beta 2$ -tubulin isoform were compiled from the literature (Popodi *et al.*, 2005; Nielsen *et al.*, 2012) and compared to the putative diamondback moth sequence

The gene has a distinctive EATA amino acid sequence as part of this tail region. The exception is the Lepidoptera with DATA instead, where glutamic acid is replaced by aspartic acid (Figure 27).

| | 1 | 10 | 21 |
|------------------|------------------|-------------------|----|
| Ae | EATADEEGE | FDEEEEGGEE-- | |
| Bm | DATAADDEGE | FDEEAEEGLEE- | |
| Dbm | DATAEEEAGE | FDEEEEAGGDEGD | |
| Ag | EATADDEGE | MDEEEEGGED-- | |
| Dm | EATADEEGE | FDEDEEGGGDE- | |
| Consensus | EATADEEGE | FDEEEEGGEE | |

Figure 27. Alignment of the $\beta 2$ -tubulin axoneme motif at the -C terminal of the gene. The position of the amino acid is given with the consensus sequence along the bottom. *Aedes aegypti*, Ae; *Bombyx mori*, Bm; *Plutella xylostella*, Dbm; *Anopheles gambiae*, Ag; *Drosophila melanogaster*, Dm.

Testis-specific transcription sets our tubulin of interest apart, with $\beta 2$ -tubulin usually the only testis-specific tubulin in a given species (Mita *et al.*, 1995). In some species the testis-specific β tubulin is not uniquely $\beta 2$, with $\beta 4$ tubulin in *Bombyx mori* also showing testis-specific expression (Kawasaki *et al.*, 1993). Recently $\beta 4$ tubulin has been used to drive testis-specific expression in *Bombyx mori*, although the gene may not be sperm-specific (Xu *et al.*, 2014).

4.1.4.3 B2-TUBULIN EXPRESSION REGULATION

Gene expression in eukaryotes is typically regulated at the transcription stage. However, the germline faces unique problems concerning gene expression during spermatogenesis compared to somatic cells with transcriptional differences relatively understudied (Sassone-Corsi, 2002). A precise stage-specific sequence of gene expression is required to differentiate from spermatogonia to spermatid to spermatozoa (Figure 28). Concurrent with this, the DNA-binding histones must be replaced with protamines. This nucleosome conversion prevents access to DNA by transcriptional machinery due to increasing chromosome compaction blocking most gene transcription (White-Cooper, 2010). This hiatus of transcription is known as meiotic arrest.

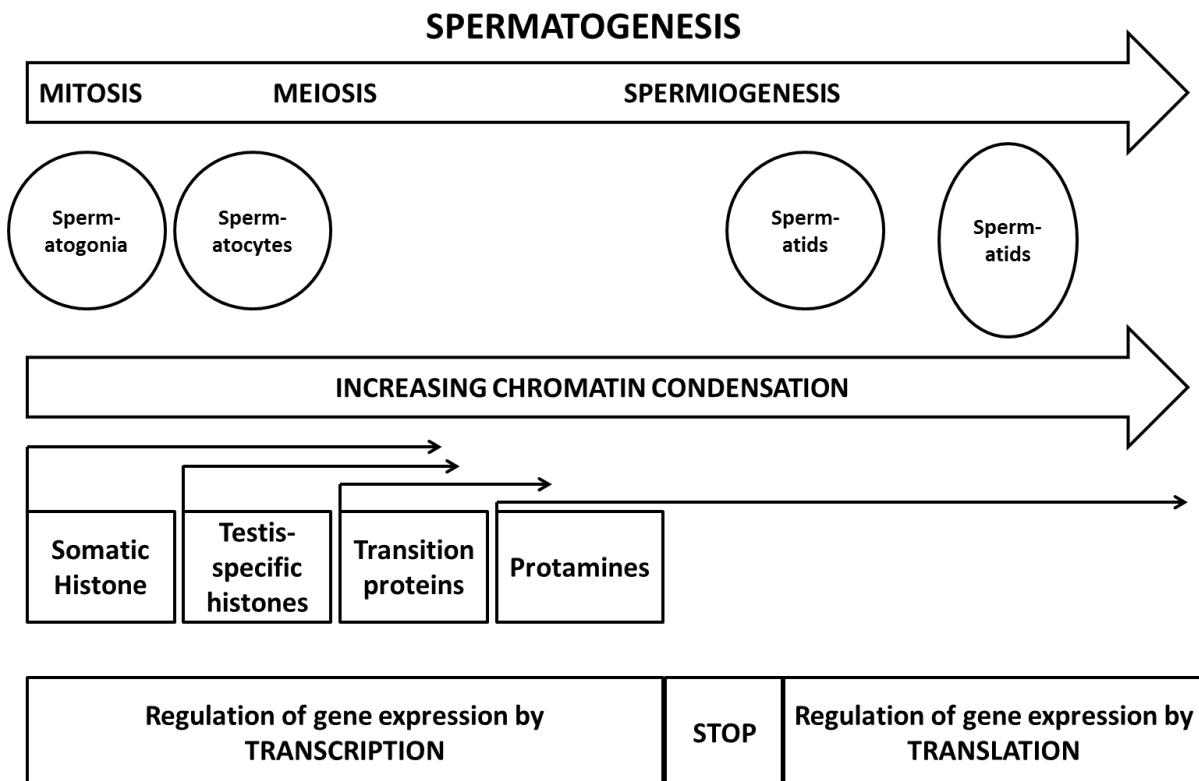


Figure 28. Spermatogenesis is a cyclic developmental process by which spermatogonia cells generate the mature spermatozoon, from left to right. These events are characterized by important modifications in chromatin organization, basically during two periods, meiosis—which includes the synapsis and desynapsis of the chromosomes—and the histone-protamine transition. Post-meiotically, a powerful wave of transcription occurs in haploid cells, which is governed by highly specialized molecular mechanisms followed by a meiotic arrest or stop in transcription. Specific genes operate at distinct steps of the spermatogenic process. After Sassone-Corsi 2002.

Meiotic arrest occurs during a period of intense protein demand due to the myriad cellular structures required to complete spermiogenesis. Therefore all (or most) transcripts must be synthesised prior to the meiotic arrest and await translation when and where required (Figure 29).

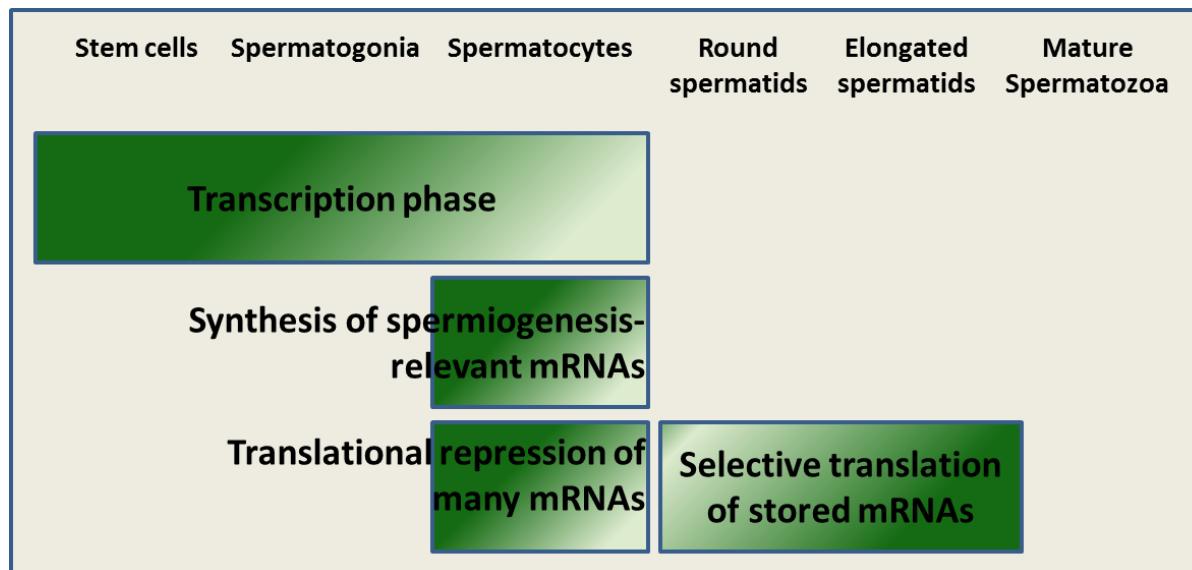


Figure 29. Spermatogenesis is characterised by the mitotic phase, meiotic phase and post-meiotic or spermiogenesis stage. In *Drosophila*, meiotic occurs after meiosis and is noticeable by hardly any transcription occurring afterwards. Instead a reservoir of hitherto translationally repressed mRNAs are used for spermiogenesis. After Rafke *et al.*, 2014.

Typically it is the spermatocytes, one of the largest cells in the male insect, which produce all the transcripts necessary for the three distinct time points in spermatogenesis (Kanippayoor *et al.*, 2013):

1. transcripts that facilitate the expansive growth during the immature primary spermatocytes stage;
2. transcripts that allow for meiosis, where transcription effectively ceases; and
3. transcripts that are translationally repressed until later post-meiotic use.

Gene expression patterns in a testis can be detected using whole-mount RNA *in situ* hybridisation (ISH), which has been used successfully to characterise gene expression in *Drosophila melanogaster* (Morris *et al.*, 2009; Zhao *et al.*, 2010). An examination of transcript distribution in a testis yields information regarding temporal and cell-specific regulation of both transcription and post-transcriptional RNA regulation.

Gene expression studies in the *Drosophila* testis have the advantage of a correlation between the position of the germ cell in the testis tube and its developmental stage. However, the method has utility in other species with gene expression patterns inferable from other data. For

example, the $\beta 2$ -tubulin gene expression pattern in a Medfly testis was investigated in the thesis of Jin (2011, The University of Oxford).

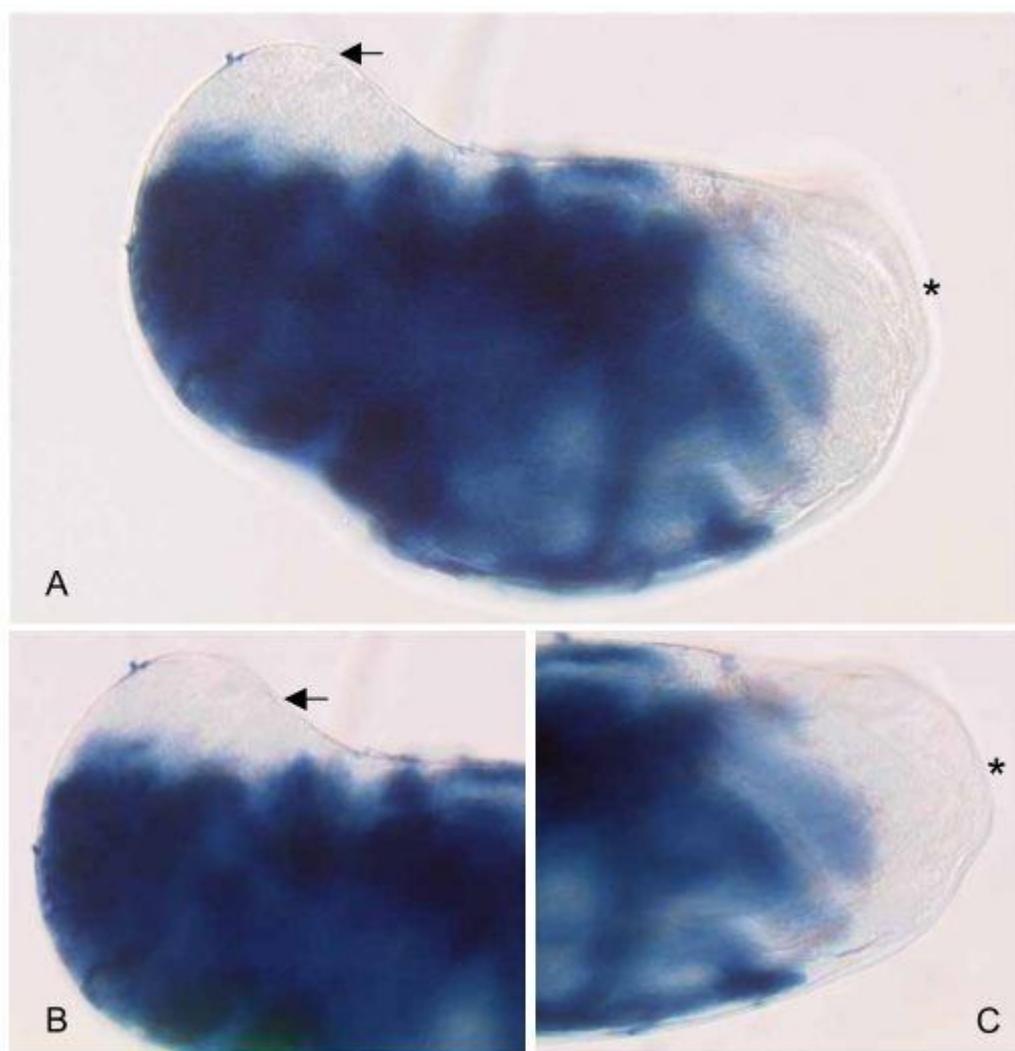


Figure 30. *Ceratitis capitata* testis RNA *in situ* hybridisation result of Cc $\beta 2$ -tubulin gene probe (the whole dark blue area shows the *Ceratitis capitata* $\beta 2$ -tubulin gene transcript location). A shows a *Ceratitis capitata* testis, the left side is the apical region of the testis (small arrow); right is distal region of the testis (*); B is a higher magnification of the apical region of the testis; C is a higher magnification of the distal region of the testis.

The ISH patterning combined with gene expression analysis in other non-testis tissues using reverse transcriptase PCR (RT-PCR) identifies candidate genes for the paternal effect system. The promoter region of the gene can then be cloned and inserted into the species of interest regulating fluorescent protein marker expression using *piggyBac*. This approach has been used both in Medfly and *Aedes aegypti* (Jin, 2011; Bilski, 2012; Catteruccia *et al.*, 2005; Scolari *et al.*, 2008; Zimowska *et al.*, 2009).

4.1.4.3.1 TRANSCRIPTION

Transcription of $\beta 2$ -tubulin in Diptera is well-characterised (Santel *et al.*, 2000), occurring at the primary spermatocyte stage. In some species, however, some non-coding regions have been necessary for transcription in addition to the promoter region, including the intron and 3' UTR (Saussede-aim & Dumontet, 2009). In *Drosophila melanogaster* it has been shown that genomic flanking regions such as the promoter, the 5' UTR and the 3' UTR can all affect the tissue specificity and expression levels of the endogenous gene (Hoyle *et al.*, 1995).

4.1.4.3.2 TRANSLATION

In *Drosophila melanogaster* translation is concurrent with transcription and occurs later during spermiogenesis (Jin, 2011). Translation of mRNA is likely tightly regulated and the control mechanism may not always be obvious, with evidence of microRNAs and other non-coding sequences playing a role in gene regulation in general (Wang & Xu, 2014).

4.1.5 IDIOSYNCRASIES OF SPERM IN THE LEPIDOPTERA

Males of most lepidopteran species produce two sperm types: eupyrene and apyrene (Carriere *et al.*, 2009). Sperm heteromorphism is thought to contribute to mating competitiveness success against other males. Each sperm type has a distinct transcriptome and regulatory elements, which may have an impact on a sperm-lethal RIDL approach in the Lepidoptera, affecting choice of regulatory sequences or resulting in interaction with the RIDL system. This may also impact the regulation of the genes we intend to use as they may be regulated differently depending on the sperm type they are expressed in.

4.1.6 RESTRICTION ENZYMES

The objectives of a sperm-lethal construct will be to introduce random double-stranded breaks (DSBs) into the nuclear DNA of the sperm at an appropriate time so as not to affect sperm

motility and competitiveness. Accordingly a short restriction enzyme recognition site will be preferred ensuring a high number of breaks per unit length of DNA. This would mitigate the risk of the evolution of resistance (Alphey, 2011). Furthermore cells have metabolic pathways for repairing double-stranded breaks. For the paternal effect system to operate, DSBs must be introduced to the DNA at a rate faster than repair mechanisms can operate.

Based on previous work with fruit flies (M. Koukidou, personal communication; Kaczorowski, 1989), *FokI* was selected as the nuclease for use in a sperm-lethal system in moths. Importantly, zinc finger nucleases (ZFN) like *FokI* require homodimerisation for activation. Furthermore, *FokI* will only dimerise at concentrations of at least 15 µM and when bound to its specific target (Miller *et al.*, 2007). This feature provides protection in the event of leaky expression in the 'off' setting, with the effect only seen when enzyme concentrations are higher.

4.1.7 CHAPERONE

The timing and level of nuclease gene expression will be determined by the promoter. The enzyme must be exposed to the DNA of the sperm cell in order to cut it. This could be achieved by coupling it with an appropriate molecular chaperone.

Protamine is a well-studied and feasible option for use as a chaperone. Protamines are chromatin-binding proteins that serve to condense and organise the DNA within sperm heads (Kanippayoor *et al.*, 2013; Rathke *et al.*, 2014). If an enzyme were expressed and fused to a protamine protein (which are very small and would not likely affect protein function, ~146 amino acids in length in *Drosophila melanogaster*) then the fusion protein would be directed to the nucleus and the DNA therein.

The protamine codon sequence is highly variable between species (Kasinsky *et al.*, 2011). The desired functionality of the protamine for the paternal effect system has been successful when using *Aedes aegypti* protamine inserted in the *Aedes aegypti* genome (Bilski, 2012). When a similar construct was inserted but with the *Drosophila melanogaster* protamine sequence

instead, the desired functionality was lost. This suggests that utilising the endogenous protamine for the paternal effect system may be important to achieve success.

4.2 HYPOTHESES

The chapter aim was to engineer a repressible male-sterility-phenotype in diamondback moth (as part of a sperm-lethal system, also known as paternal effect). This can be broken down into hypotheses, bearing in mind that for diamondback moth, the sex sorting and conditional modules were already in place.

- The orthologue for $\beta 2$ -tubulin can be found in diamondback moth, sequenced and confirmed by RT-PCR and *in situ* hybridisation of the testes to show testis specificity.
- $\beta 2$ -tubulin promoter region in diamondback moth can be identified and provide (male) sperm-specific transcription.
- A fluorescent protein linked to the $\beta 2$ -tubulin promoter will result in fluorescent sperm and/or testes in late larvae and/or young adults.
- A suitable protamine-like gene can be identified in diamondback moth for use in the paternal effect system.
- A “paternal effect system” will provide effective sterility phenotype in male diamondback moth using the *Fok1* nuclease and the components given above.

4.3 RESULTS AND DISCUSSION

4.3.1 DETERMINATION OF GENE EXPRESSION PATTERNS OF PATERNAL EFFECT CANDIDATE GENES IN DIAMONDBACK MOTH TESTES USING *IN SITU* HYBRIDISATION

4.3.1.1 INTRODUCTION

4.3.1.1.1 TESTIS ANATOMY AND DEVELOPMENT IN DIAMONDBACK MOTH

Dissection of the testes at different life stages elucidated the changes associated with sexual maturation in the male (Figure 31).

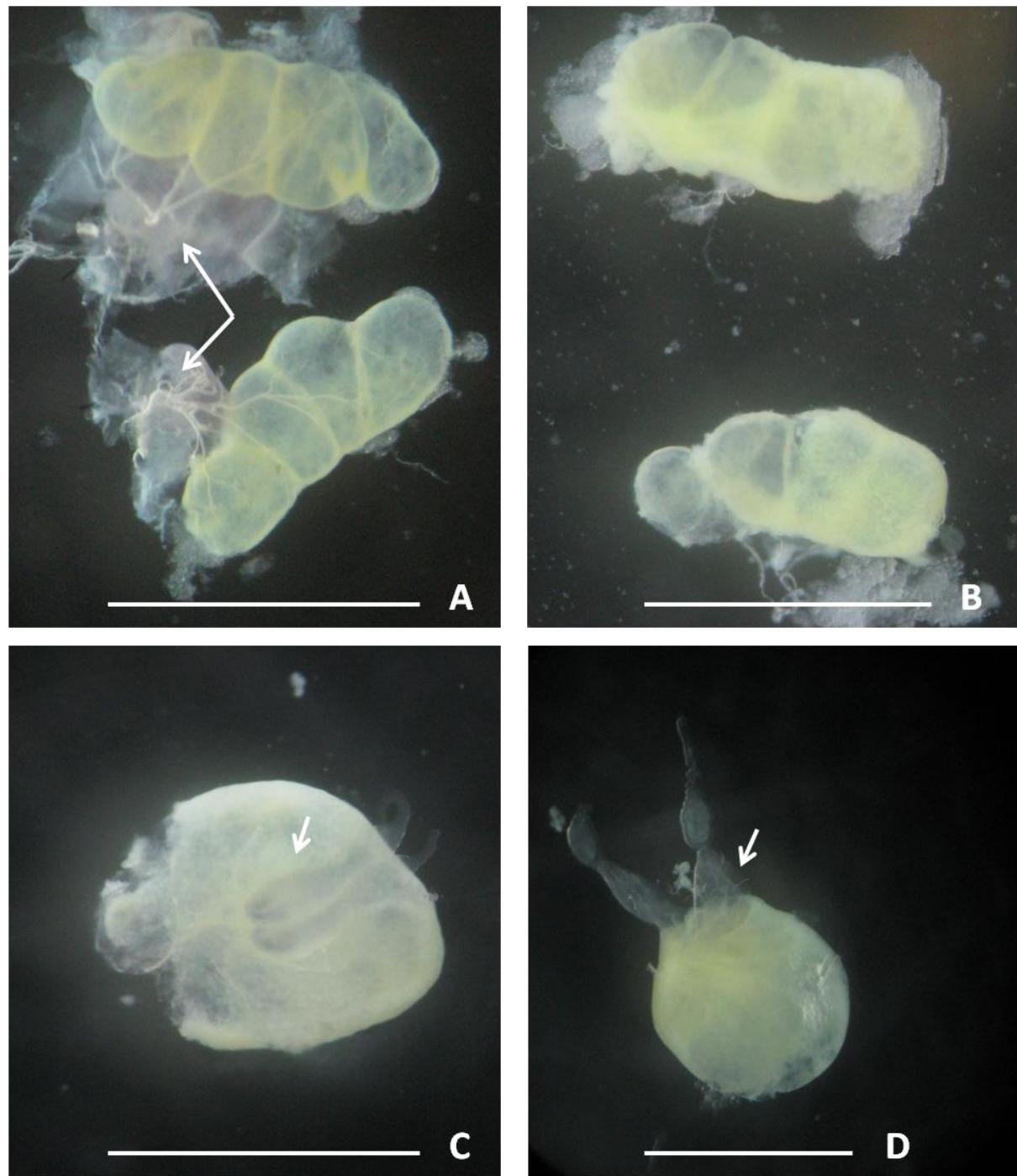


Figure 31. Testes of *Plutella xylostella* at different post-embryonic stages. Observed in a stereoscopic microscope: (A) 10-day old larva; arrows indicate trachea from the fifth abdominal segment, (B) Pre-pupa; follicles are more compact, (C) Pupa; testes are fused with vas deferens (arrowed) where sperm are stored, (D) Adult; spherical and more opaque. The bar represents 1mm. The vas deferens are more obvious in C and D.

Dissection and microscopy revealed anatomical and morphometric changes to the testes in diamondback moth through post-embryonic development from late larva to adult (Table 9) not well characterised in the literature. The four follicles per testis are visible in all life stages, with inter-follicle fusion at the pupal stage.

Table 9. An observational study of diamondback moth testes morphological changes through larva to adulthood.

| Life stage | Testes shape | Colour | Follicle status | Testes |
|------------------------------|--------------|----------------------|------------------------------|-----------|
| 4 th instar larva | Reniform | Yellow / transparent | Individual follicles visible | Not fused |
| pupa | Oval | White / opaque | Barely visible | Fused |
| 2-day old adult | Spherical | Yellow / opaque | Not visible | Fused |

The anatomical aspects of diamondback moth conform to the general trend seen in the Lepidoptera.

4.3.1.2 COLLABORATIVE WORK SUMMARY

Colleagues at the Centre for Genomic Research, Liverpool, had provided us with diamondback moth genome and RNA-seq transcriptome data (detailed in Chapter 1). The data provided us with a starting point and information regarding relative transcript intensity in different diamondback moth tissues including male carcass, female carcass, male testes and female ovaries based on our wild-type diamondback moth strain. The genes with relatively high and specific transcription in the male testes were compiled into a list of top 100 candidate genes (where the first had the highest relative testis-specific transcription to carcass ratio). These genes were filtered for those with plausible fruit fly orthologues by BLASTing against Flybase (<http://flybase.org/>) and FlyTED (currently offline); leaving 20 candidates (Bioinformatics

work conducted by Hongfei Gong; I was involved with the manual sorting). The genes were labelled according to their fly orthologues and by their rank for top testes transcription.

4.3.1.3 CANDIDATE GENES FOR PROBE DESIGN

I selected five genes from the top-20 list to analyse by RT-PCR (Table 10) to validate the transcriptome data. The selected genes were chosen as they had supporting evidence from FlyAtlas or FlyTED that they had testis-enriched expression in *D. melanogaster* (Table 10).

Table 10. A summary table of the evaluation of five genes from the top twenty list of the RNA seq Liverpool diamondback moth data for transcript relative quantity in testes. The gene sequences from the diamondback moth scaffolds were BLASTed against FlyBASE and FlyTED. Supporting evidence for testis-specific transcription was positive if blue-filled in the FlyATLAS column and or orange-filled in the FlyTED column. The top 20 list was compiled by HongFei Gong and checked by myself, it was then narrowed down to the five candidates described here.

| Order of genes In diamondback moth Top 20 list | ID to gene name Conversion | | Hits in gene expression database | | |
|--|-----------------------------|--------------------|-------------------------------------|---------------------|--------|
| | Submitted ID | Related record | FlyAtlas | Testis specific? | FlyTED |
| 1 | FBpp0081495 | CG8121-PA | | | Yes |
| | FBpp0081496 | CG8121-PC | | | Yes |
| | FBpp0081497 | CG8121-PB | | | Yes |
| | | | | | |
| 4 | FBpp0072177 | betaTub60D-PA | | | |
| | FBpp0081524 | betaTub85D-PA | | | Yes |
| | FBpp0084630 - unknown ID | | | | |
| | FBpp0085720 | betaTub56D-PB | | | |
| | FBpp0085721 | betaTub56D-PA | | | |
| | FBpp0289838 | betaTub97EF- PB | | | |
| | | | | | |
| 5 | FBpp0078832 | CG9222-PA | Yes | Yes | Yes |
| | | | | | |
| 6 | FBpp0086303 | CG15925-PA | Yes | Yes | |
| | FBpp0086625 | CG12869-PA | | | |
| | | | | | |
| 17 | FBpp0073944 | CG12698-PA | Yes | Yes | |
| | FBpp0073987 | CG8958-PA | Yes | Yes | |
| | FBpp0081777 | CG14693-PB | | | |
| | FBpp0289355 | CG14693-PC | | | |
| | FBpp0289356 | CG14693-PD | | | |

The FlyTED database provided additional information including a photograph of *in situ* hybridisation of the transcript of the gene of interest in *D. melanogaster* testes for some of the genes described above. These images will be discussed in more detail when comparing transcription patterning in diamondback moth.

The gDNA sequences were assessed and RT-PCR primers designed to confirm testes-specific expression (Figure 32). RNA was extracted from the diamondback moth fourth instar larvae testes and compared to male (upper part of carcass, no testes) and female carcasses by RT-PCR with several primer pairs for each gene (Methods 8.2.25).

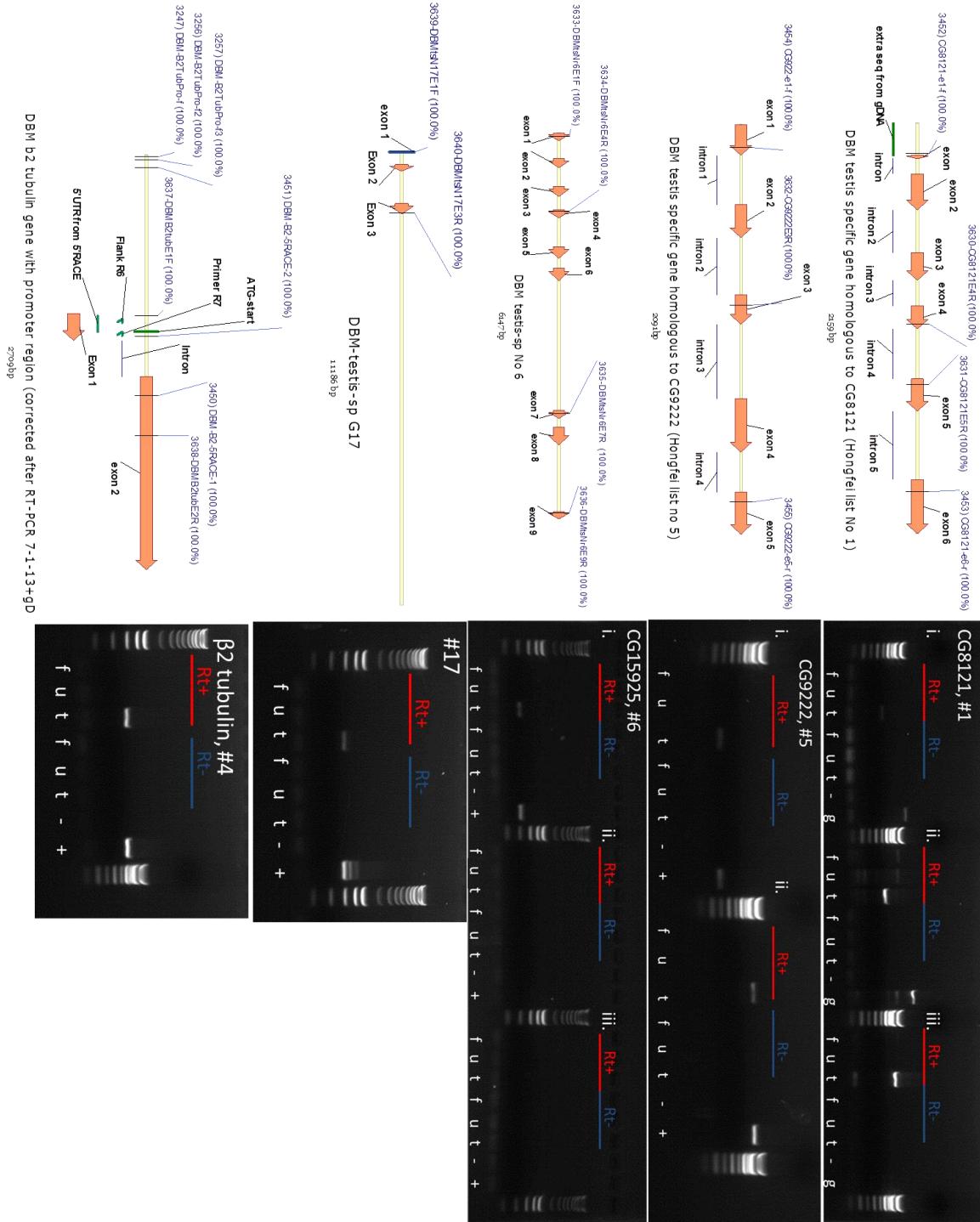


Figure 32. Genetic feature maps of putative testis-specific genes of interest in diamondback moth including designed primers and targets of amplification in RT-PCR (gene maps on the left and RT-PCR outcomes for adjacent gene on the right). The putative exons and introns were elucidated by gDNA and primary transcript mRNA comparison using Vector NTI. Primer positions are provided with sequence similarity in brackets. The putative gene identity and size are provided at the bottom of the gene map. RT-PCR of OX4804 (lines C, D and E) derived RNA from three sample types; female, f, upper half of fourth-instar male larva excluding testes, u and lower half of fourth-instar male larva including testes, t. Samples were either exposed to reverse transcriptase (Rt+) or not (Rt-) to produce cDNA (or not) for use in a PCR. For #1, three primer pairs were used, i) 477 bp (gDNA 921 bp), ii) 528 bp (gDNA 1235 bp) and iii) 734 bp (gDNA 1802 bp). For #5, Two different reactions were conducted using primer pairs with estimated PCR products for cDNA amplification at i) 282 bp and ii) 706 bp, for #6 three different reactions were conducted with estimated PCR products for

Chapter 4 | Engineering male sterility in diamondback moth

cDNA amplification at i) 466 bp, ii) 1021 bp and iii) 1401 bp, for #17 one reaction was conducted estimated PCR products for cDNA amplification at i) 472 bp, for #4 the gene sequence shared similarity with the *Drosophila melanogaster* gene $\beta 2$ -tubulin gene with estimated size of 574 bp. Primer pairs selected for probe design in RISH are circled if a preference was required. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

The RT-PCR validated the Liverpool data suggesting testis-enriched expression of the selected genes. I decided to use whole-mount RNA *in situ* hybridisation (ISH) to provide a semi-quantitative measure of gene expression and location in the diamondback moth testes. I applied the *D. melanogaster* protocol to diamondback moth (Methods 8.2.25 modified from Morris *et al.*, 2009) testes with some modifications and troubleshooting. The primer pairs that provided the strongest amplification were used for DIG-probe development for RISH.

4.3.2 *IN SITU* HYBRIDISATION

Prepared testes must be made amenable to probe penetration, hybridised and then washed prior to substrate addition and imaging.

4.3.2.1 TESTES PREPARATION FOR *IN SITU* HYBRIDISATION

The diameters of the pupal testes used for ISH were measured to the nearest millimetre upon dissection (Figure 33). Measuring the size allowed an estimation of the surface area-to-volume ratio compared to *D. melanogaster* (not shown), informing the protocol of the proteinase K step (step 24 in Morris *et al.*, 2009).

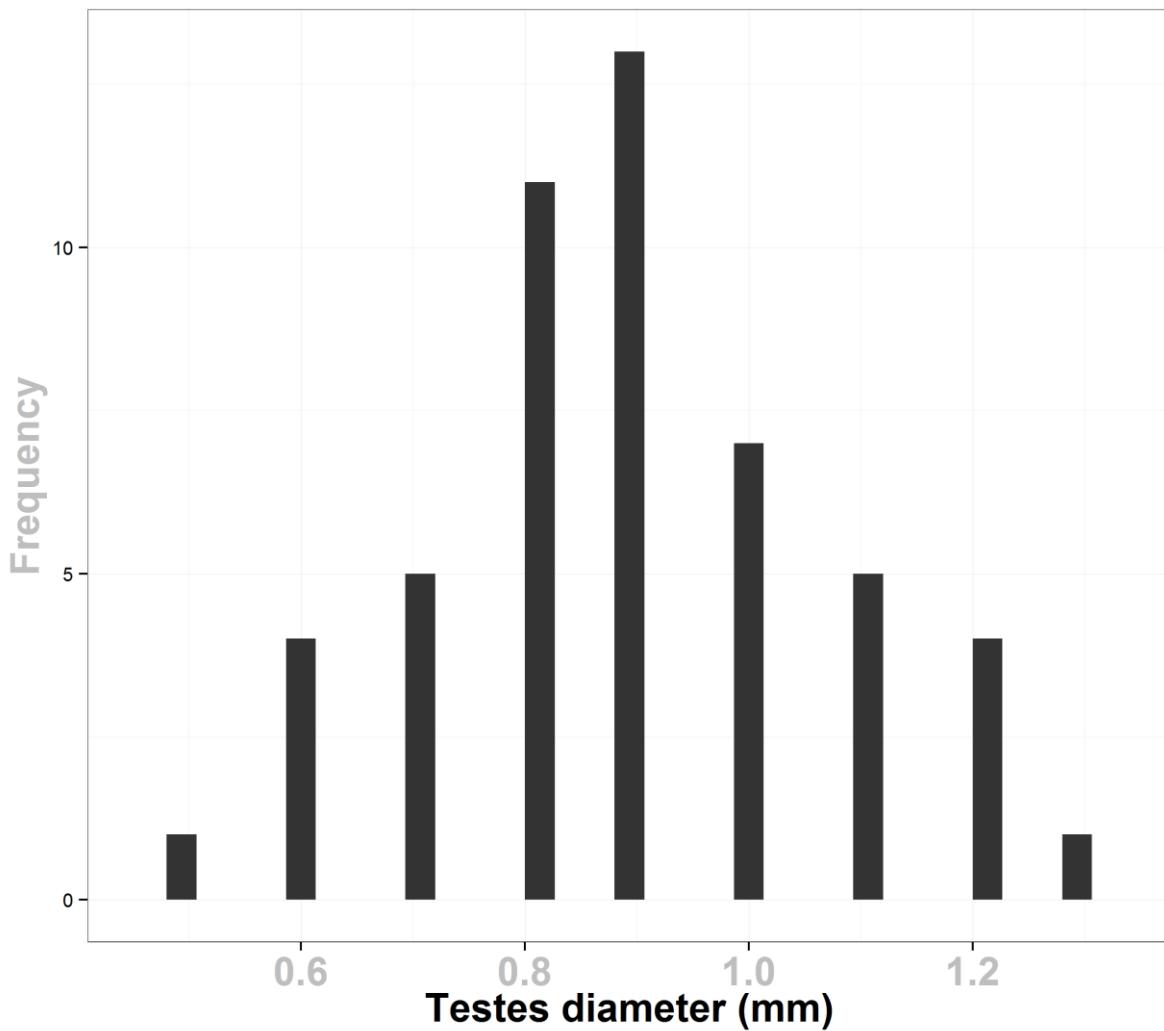


Figure 33. A histogram of the diameters of dissected pupae fused-testes (mean = 0.89, 95% CI of 0.84-0.94 mm).

4.3.2.2 VISUALISING TARGET-GENE SPECIFIC mRNA IN THE TESTES OF

DIAMONDBACK MOTH

4.3.2.2.1 #1 CG8121 - PROBES 1 & 2

The anti-sense probe (probe 1) for the diamondback moth gene homologous to CG8121 and our selected gene #1 stained the trachea and parts of the vas deferens, but were otherwise removed by the washes (Figure 34).

The anti-sense probe bound to the transcript of the gene which had been previously identified as a testis-specific gene. The ISH provides greater resolution and shows that, during

spermatogenesis, the transcript is found where the spermatids develop into spermatozoa. This was inferred by the higher magnification images at x10 and x40 where the distinctive sperm flagella are visible and the sperm bundles are relatively large. The transcript appears to be found ubiquitously within the mature spermatids and are not localised to the head or tail.

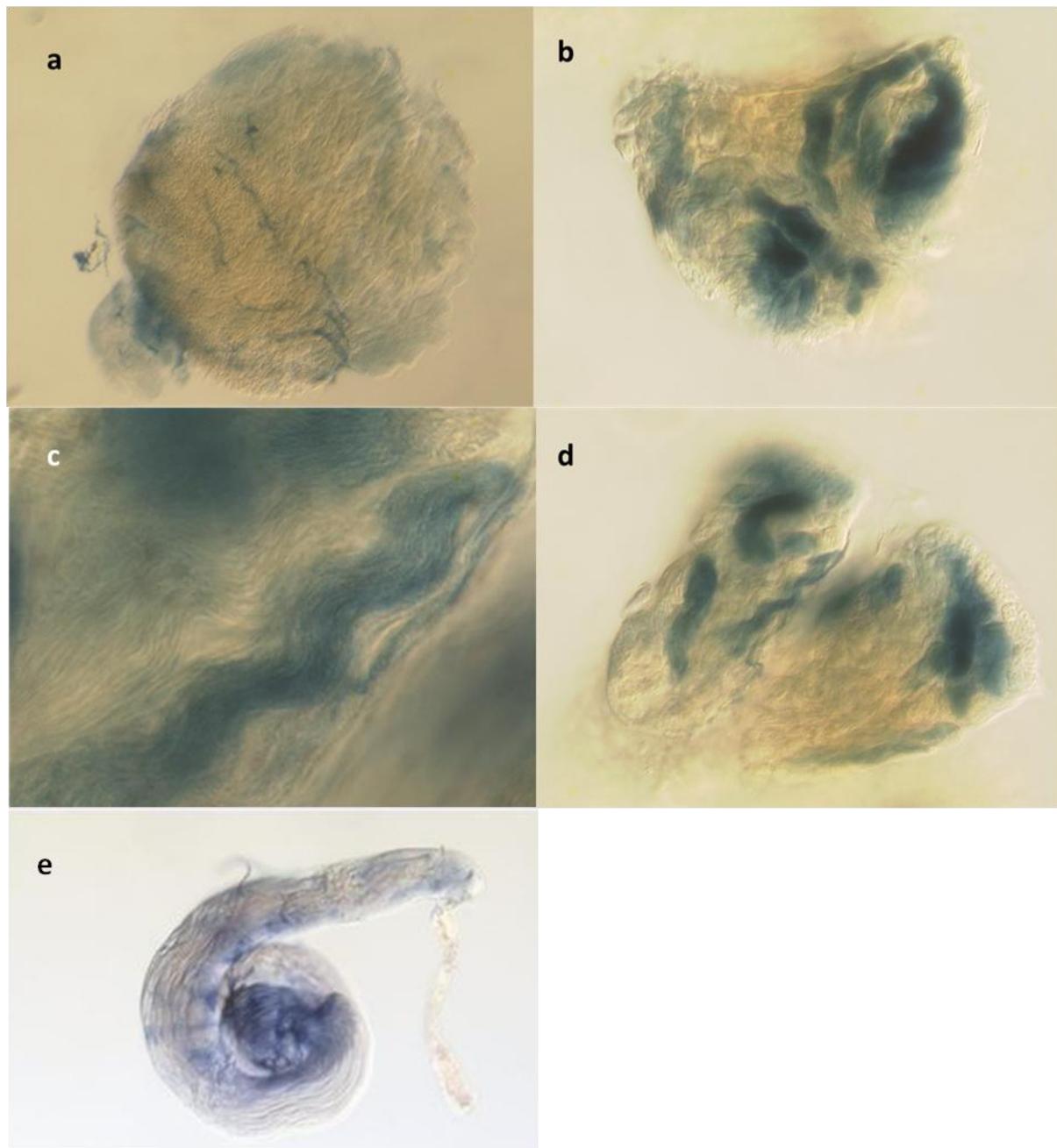


Figure 34. Diamondback moth CG8121 orthologue RNA *in situ* hybridisation in diamondback moth testis. A comparison of sense (a) and anti-sense probes (b-d). Expression pattern of CG8121 shown at x 4 magnification in c and d. The transcript was present in the post-meiotic sperm cells. The sperm bundles are stained and are visible with individual sperm flagellum also visible. The testes were left in the dark with the staining reagents for two hours. e) Microphotograph of *in situ* hybridisation the CG8121 gene in wild-type *D. melanogaster* testis. Post-meiotic gene expression in spermatocytes is evident. Taken from the FlyTED database prior to closure.

The follicular structure and testes morphology was lost and difficult to interpret following the staining process. The sperm bundles appear to be homogenously stained and in the latter stages of maturation. Utilising the testes diameter estimate of 890 µm in pupae allows us to estimate the length of the sperm bundle as approximately a third of the length of a follicle or 150 µm.

In FlyBase the gene CG8121 for which our #1 diamondback moth gene shares some sequence homology is described as a protein-coding gene but with no known molecular function either predicted or evidenced. In a FlyATLAS search, the gene showed relatively low levels of mRNA in the testis compared to other tissues. This suggests that the diamondback moth gene, putatively named CG8121 after a BLAST against Drosophila genome, may not be orthologous to this gene. This inference is based on it being testis-specific, as shown by the previous RT-PCR and also clearly sperm (or spermatid)-specific, elucidated by the ISH.

A visual comparison between CG8121 ISH in Drosophila testes compared to diamondback moth tests shows a similar expression pattern between them. In *D. melanogaster*, this gene is expressed in early spermatocytes and is stable to late primary spermatocytes with some evidence of late expression in maturing sperm (Figure 34e). The sperm bundles are not as clearly defined as in diamondback moth.

The late transcription in spermiogenesis of gene #1 in diamondback moth renders it inappropriate for the paternal effect system.

4.3.2.2.2 #5 CG9222 - PROBES 3 & 4

The sense probe stained the trachea and outer edges of the testes. Some non-specific staining is present throughout, suggesting an over-long staining process. The anti-sense probes show high levels of the transcript, which was distributed evenly in the follicles (Figure 35b and d). A higher magnification reveals some patchiness and background noise; however no clear pattern was discernible other than the transcript being widespread except in control tissue, such as the vas deferens. The ISH process left some of the testes slightly degraded and no longer fused.

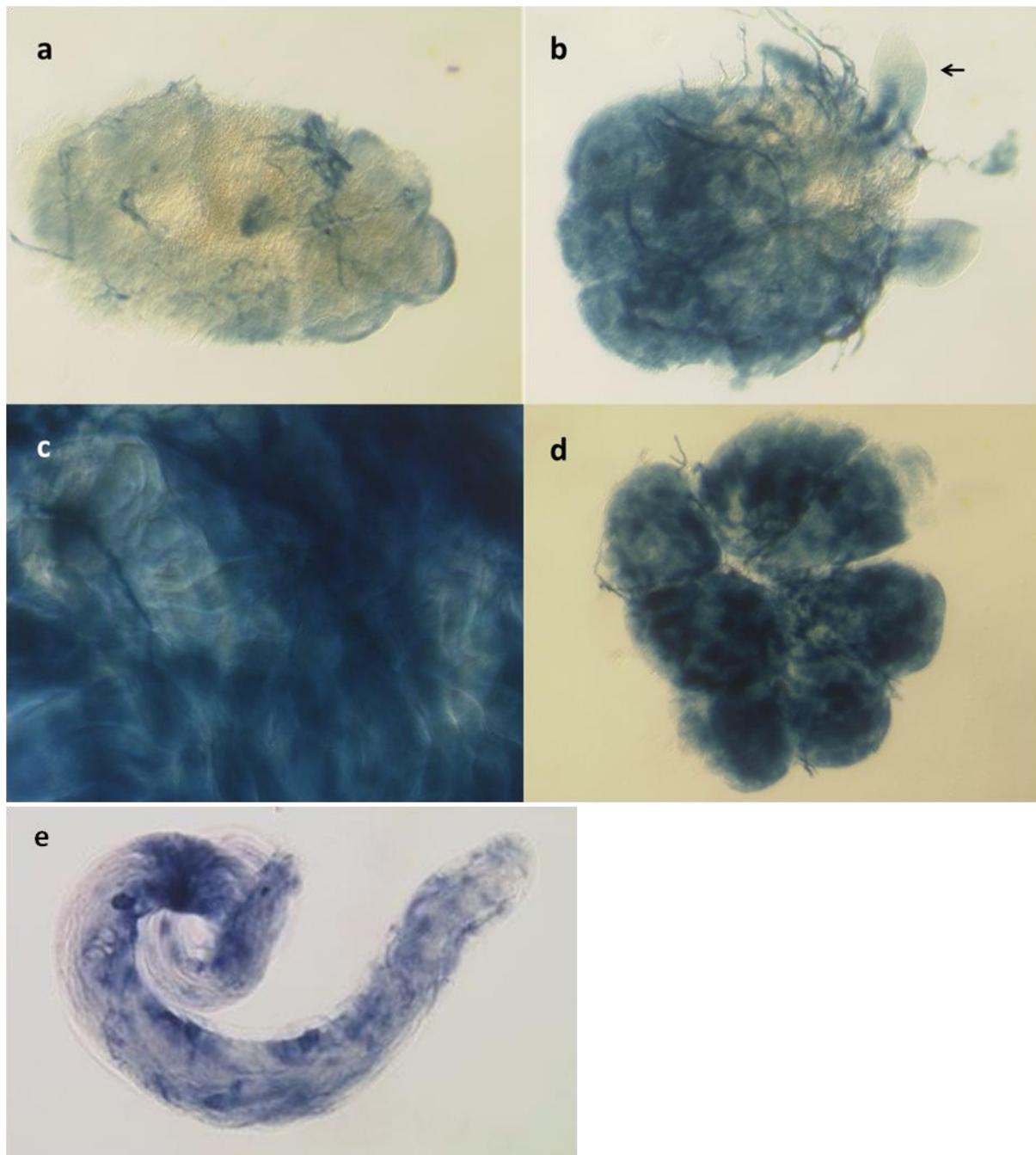


Figure 35. Diamondback moth CG9222 orthologue RNA *in situ* hybridisation in diamondback moth testis. A comparison of sense (a) and anti-sense probes (b-d). Expression pattern of CG9222 shown at x4 magnification in c and d. The transcript was present throughout the testes without any discernible pattern. Sperm were not visibly stained. The testes were left in the dark with the staining reagents for 2 h. The black arrow indicates the vas deferens. e) Microphotograph of *in situ* hybridisation the CG9222 gene in wild-type *D. melanogaster* testis from the *Drosophila* testis gene expression database. Cyst cell staining is evident from the darkened cells.

The spatially ubiquitous transcription of CG9222 suggests that it has a gene product required at all stages of spermatogenesis. Summary data from FlyBase (<http://flybase.org/reports/FBgn0031784.html>) describes this gene as having a protein-kinase

domain and protein-kinase activity. Expression levels are high in adult testis and moderate in larval fat body as shown by a FlyATLAS search. The presence in the larval fat body could be due to contamination of the testes in the fat body — although moderate mRNA signals were in 4 of 4 replicates — due to the difficulty in a clean dissection of fat bodies (testes contamination) in larvae. Peak expression is observed in adult male testis with relative levels of mRNA ten times that found in the fat body (Liverpool transcriptome data).

The ISH of the diamondback moth homologue for this gene suggest ubiquitous transcription throughout the testes with other life stages' testes requiring examination to permit comparison. Protein kinases are key regulators of cell function and the diamondback moth orthologue of CG9222 may be involved in spermatogenesis regulation.

The cognate gene in *D. melanogaster* testes showed a similar expression pattern except for an obvious high concentration of the transcript in the cyst cells (Figure 35e). This gene is probably unsuitable for the paternal effect system.

4.3.2.2.3 #6 - PROBES 5 & 6

The staining showed high background staining with both probe types (not shown). Due to the failure of the negative control (sense probe), inferences from the anti-sense probe ISH were rendered unreliable.

4.3.2.2.4 #17 - PROBES 7 & 8

The staining showed high background staining with both probe types (not shown). Due to the failure of the negative control, inferences from the anti-sense probe ISH were rendered unreliable.

4.3.2.2.5 #4 B2-TUBULIN - PROBES 9 & 10

The sense probe images were particularly clean with no background staining of trachea or vas deferens evident for probe 9. This may have been due to this probe being left for the shortest

staining period of 1 h, although some staining of trachea was evident in the anti-sense probe images.

The anti-sense probe identified the transcript relative location in the testes, with the pattern indicating high transcript concentration in the follicles with a declining gradient in transcript concentration towards the interior of the testes, or towards the vas deferens. This suggests that the transcript is present in the early and middle stages of spermatogenesis and is translated for spermiogenesis and associated flagella development (Figure 36).

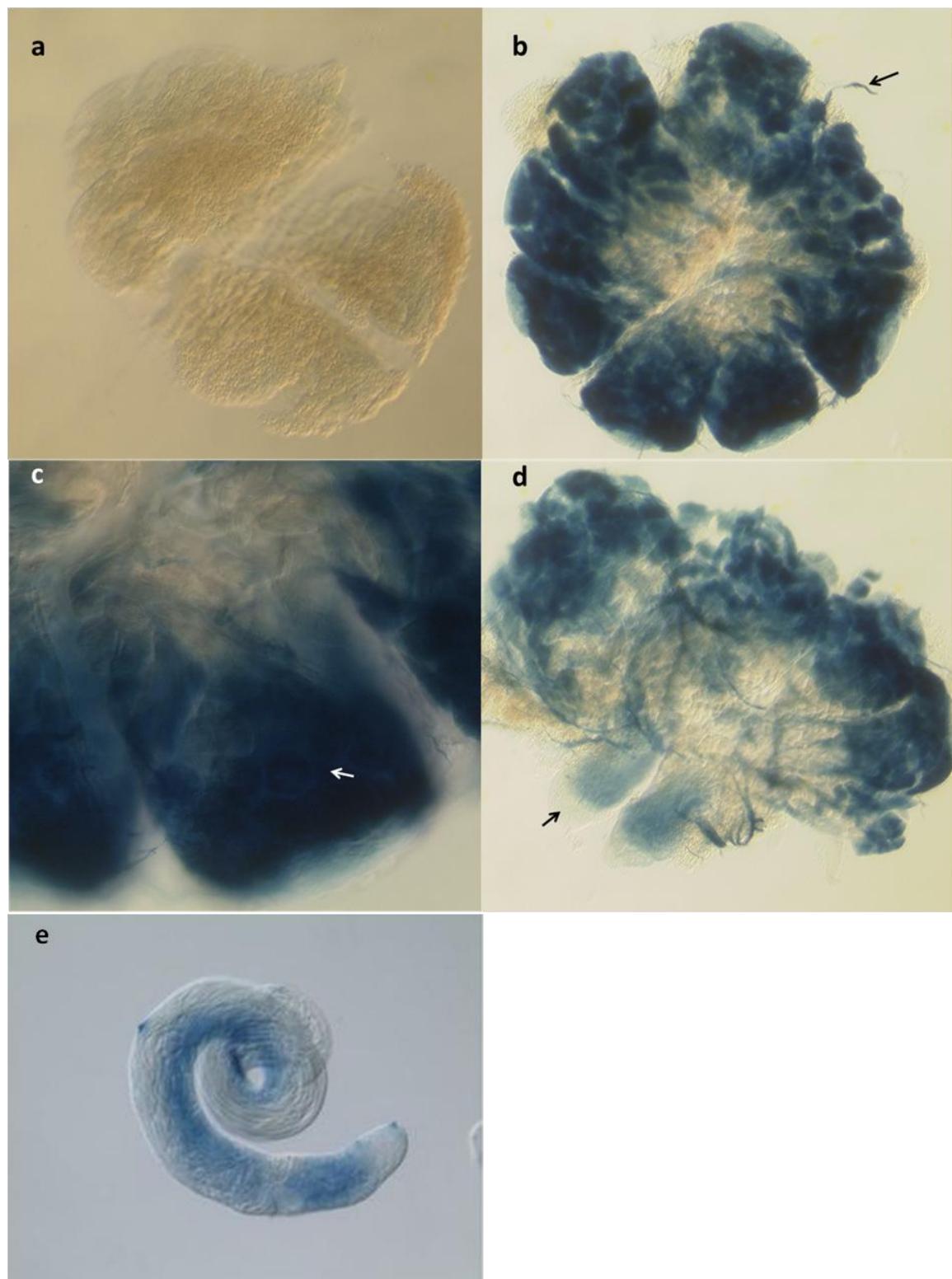


Figure 36. Diamondback moth $\beta 2$ -tubulin orthologue RNA *in situ* hybridisation in diamondback moth testis. A comparison of sense (a) and anti-sense probes (b-d). Expression pattern of $\beta 2$ -tubulin shown at x 4 magnification in b and d with the black arrow showing probable vas deferens location. The testes were left in the dark with the staining reagents for one hour. c shows a close up of a follicle at x10 magnification circular cysts may be visible and are shown with a white arrow. e shows a microphotograph of *in situ* hybridisation of the *Drosophila melanogaster* $\beta 2$ -tubulin gene in wild-type *D. melanogaster* testis from the *Drosophila* testis gene expression database.

Summary data from FlyBase describes this gene saliently as having a tubulin domain and a cytoskeleton function. A FlyATLAS search (gene code CG9359) indicated high relative levels of mRNA in the testis.

A similar gradient of early or apical transcript location to reduced transcript at the distal end was observed in Medfly testes for the *Ccβ2*-tubulin gene. Transcription is thought to occur post-mitotically and prior to meiosis (Figure 37). Comparison with *D. melanogaster* revealed a similar expression pattern, with the gene expressed in mid-to-early spermatocytes, stable-to-late primary spermatocytes, and persisting to early elongation. The staining is not as pronounced in *D. melanogaster* but follows a similar pattern (Figure 36e). With *β2-tubulin* being well conserved between species, it is perhaps unsurprising that the expression pattern and function of this structural gene remains similar in insects.

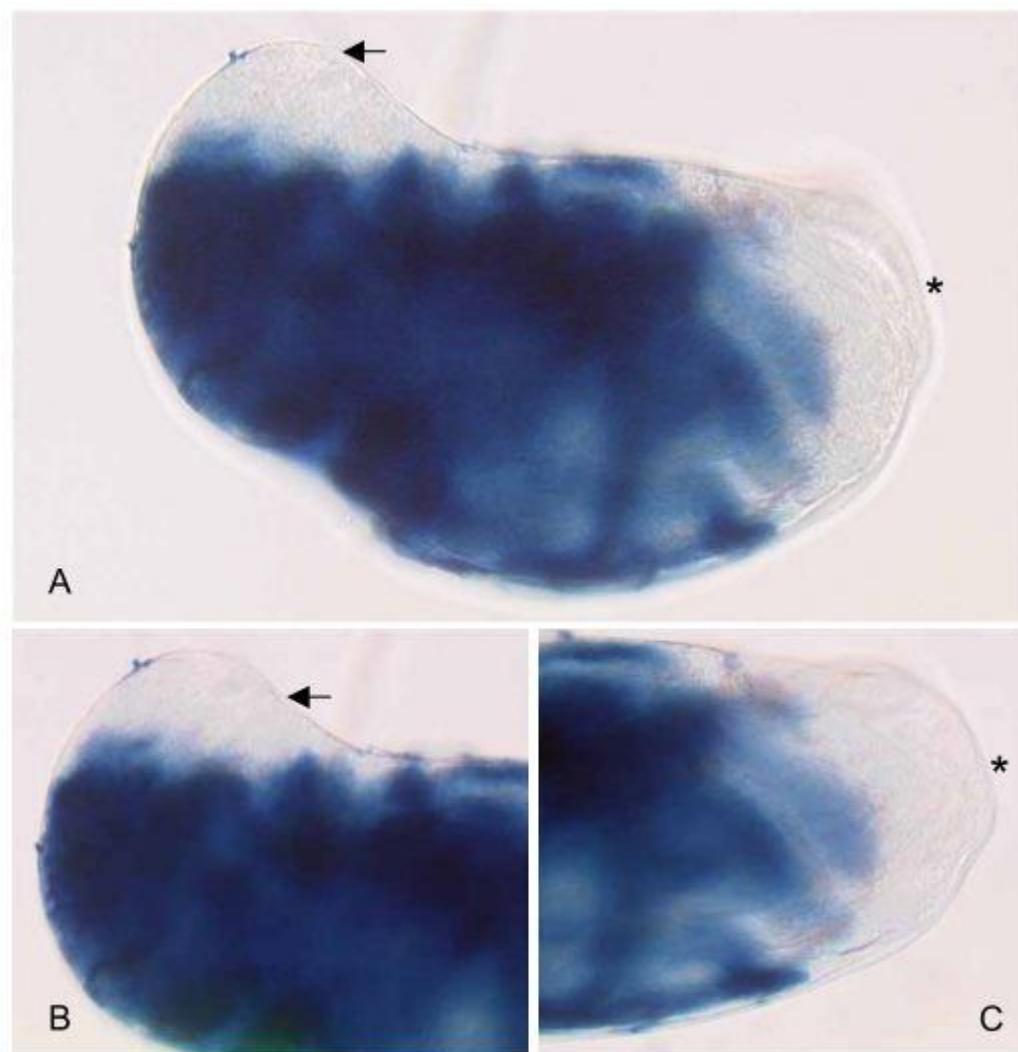


Figure 37. Medfly testis RISH result of *Cc β 2-tubulin* gene probe (the whole dark blue area shows the Medfly β 2-tubulin gene transcript location). (A) shows a Medfly testis, the left side is the apical region of the testis (small arrow); right is distal region of the testis (*); (B) is a higher magnification of the apical region of the testis; (C) is a higher magnification of the distal region of the testis (After Jin, 2011).

Of the five gene expression patterns examined only three were successfully probed with ISH. β 2-tubulin demonstrated a potentially viable expression pattern as part of a paternal effect system.

4.3.3 TOWARD A PATERNAL EFFECT SYSTEM IN DIAMONDBACK MOTH

Based on recent success in developing a paternal effect system in Medfly, it was decided to use a similar design of construct intended to express DsRed2 in the nucleus of the sperm (Jin, 2011).

The DsRed2 would be chaperoned to the nucleus by protamine derived from *D. melanogaster* regulated by the diamondback moth β 2-tubulin promoter.

4.3.3.1 OX4703- ATTEMPTED SPERM-MARKING IN DIAMONDBACK MOTH

A construct was designed for insertion into the germline of diamondback moth, with transformants marked with AmCyan regulated by the Hr5-*ie1* enhancer-promoter, which has previously been shown as a strong regulator of marker expression in transformed diamondback moth (Martins *et al.*, 2012). OX4703 was designed to test the function of the putative *β2-tubulin* promoter from diamondback moth as a regulator of sperm-specific expression, as indicated by the fluorescent protein, DsRed2. Testis-specific genes typically have short promoter regions (White-Cooper *et al.*, 2010) hence a non-excessive amount of the upstream region of the *β2-tubulin* gene (955 bp) was included in the construct. The 5' UTR was presumed to be important for testis specificity so was included.

The inclusion of protamine from *D. melanogaster* was intended to function as a molecular chaperone for the DsRed2; transporting it into the sperm nuclei. Successful insertion would be indicated by Hr5-*ie1*-AmCyan expression patterning and DsRed2 in the sperm nuclei (Figure 38). The 5' UTR of the protamine was included and followed that of the *β2-tubulin*. It was assumed that these would be ignored until the start codon in the protamine. Translation of this small protein would be fused to the larger DsRed2. The AmCyan was fused to nuclear localisation signal (nls) (at both ends) to assist in getting the marker into the nucleus and rendering it more visible (see Appendix 9.1.2 for design and synthesis details).

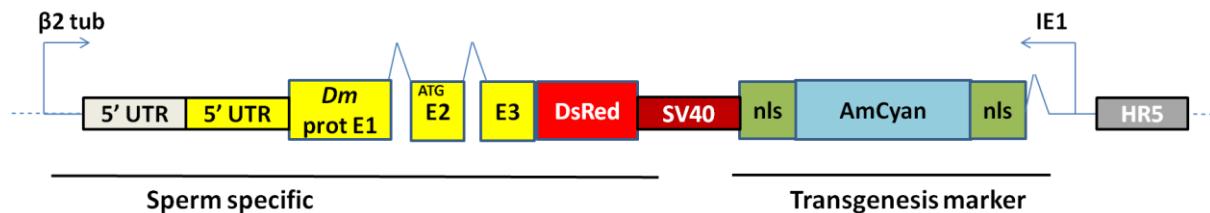


Figure 38. A simplified representation of the genetic construct OX4703. The $\beta 2$ -tubulin promoter was derived from *Plutella xylostella* genome inspection and successfully cloned. The main transgenesis marker is provided by the immediate early promoter (*Hr5-*ie1**) with AmCyan-SV40. The sperm specific components were regulated by about 1 kb upstream of diamondback moth $\beta 2$ -tubulin. The protamine was derived from *Drosophila melanogaster* (yellow) and was fused to DsRed2-SV40. The dashed line represents other components of the circular construct. Throughout this thesis genes are typically followed by SV40 3' UTR (burgundy) unless stated otherwise. The Dm protamine also includes the 5' UTR (yellow). The $\beta 2$ -tubulin includes the 5' UTR of the gene and the promoter (cream).

4.3.3.2 TRANSGENESIS

Four transgenic lines were generated from just over 200 G₀. Transgenic G₁ insects were assigned an identifying code and then crossed with a wild-type counterpart. G₂ insects were screened at pupae for the marker and then counted (Figure 39A) to assess for multiple insertions of the construct. In strains A and B there was weak evidence in favour of the null hypothesis of a 1:1 ratio, possibly explained by failure of auxiliary assumptions being met; i.e. differential survival of transgenics and wild-types due to position effects of the insertion.

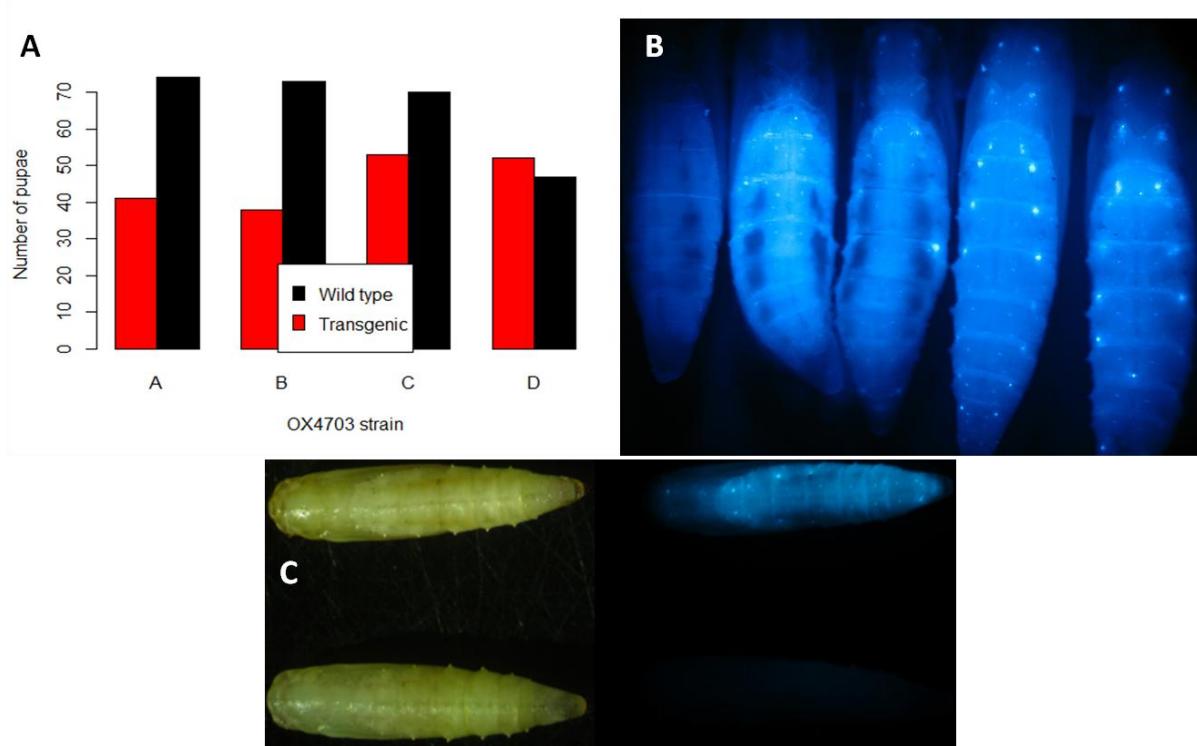


Figure 39. Transgenesis of diamondback moth with OX4703. A) OX4703 G₂ progeny status (G₁ transgenic crossed with wild-type). The numbers of transgenic and wild-type pupae derived from G₂ crosses of four strains, OX4703A, OX4703B, OX4703C and OX4703D, created from OX4703 injection into diamondback moth. A Chi-squared test was conducted on OX4703A data compared to a 1:1 ratio expected from a Mendelian punnett square of a dominant heterozygote marker gene ($\chi^2 = 4.2661$, df = 1, p = 0.03888). The same test was applied to the other lines: OX4703B ($\chi^2 = 5.0344$, df = 1, p = 0.02491), OX4703C ($\chi^2 = 0.919$, df = 1, p = 0.3377), OX4703D ($\chi^2 = 0.0414$, df = 1, p = 0.8394). B) OX4703 diamondback moth transgenic pupae. Dorsal view, from left to right: wild-type, OX4703, OX4703B, OX4703C, OX4703D. A fluorescent microscope with cyan filters was used. The marker is more apparent when viewing the insect from the dorsal side. Inter-strain variation should not be judged from this photo as intra-strain variation in fluorescent patterning was evident in all strains. The trend was an overall fluorescence with points of high expression as spots along the dorsal side. If expression differences existed it may be due to differences in insertion position in the genome (Kapetanaki *et al.*, 2002) or inter-individual variation. C) Dorsal view: wild-type bottom row, transgenic 4703A top row. A fluorescent microscope with cyan filter was used for the right hand photographs of the same insects. A white light LED ring and no filter were used for the left photographs. The marker is more apparent when viewing the insect from the dorsal side as evidenced by the localisations of fluorescent protein giving rise to the characteristic dotted appearance.

At the macroscopic level the fluorescent protein tends to aggregate along the dorsal side of the organism probably due to somatic cells with enlarged nuclei or a higher concentration of cells in that area (Figure 39B). These are visible as bright points of fluorescence and are visible at low magnifications (40×). This confirms previous work by Martins *et al.*, (2012). The Hr5-*ie1*-AmCyan shows successful transformation clearly in the insect (Figure 39C). A caveat, we would not detect those instances where it did insert, but was not detectable. These false negatives could explain the higher ratio of wild-type to transformants detected in the crosses.

Microinjection data and transformation statistics are tabulated for convenience as is typical in the literature and as discussed in Chapter 2.

Table 11. The injection data summary for the construct OX4703 in diamondback moth. Proportion survival and transformation efficiency given in brackets.

| Species | Construct | Embryo injections (n1) | Injection survivors (G ₀) | Independent transgenic lines (X) |
|----------------------------|-----------|------------------------|---------------------------------------|----------------------------------|
| <i>Plutella xylostella</i> | 4703 | 779 | 206 (0.26) | 4 (0.02) |

4.3.3.4 TESTES DISSECTION AND FLUORESCENCE MICROSCOPY

Dissection of a number of wild-type adult diamondback moth was performed in order to consistently locate the testes as some skill and experience is required to identify and remove the organ. Published information from Justus and Mitchell (1999) was used to provide guidance to identify the testes (Figure 40A). Testes were confirmed by locating the distinctive sperm bundles.

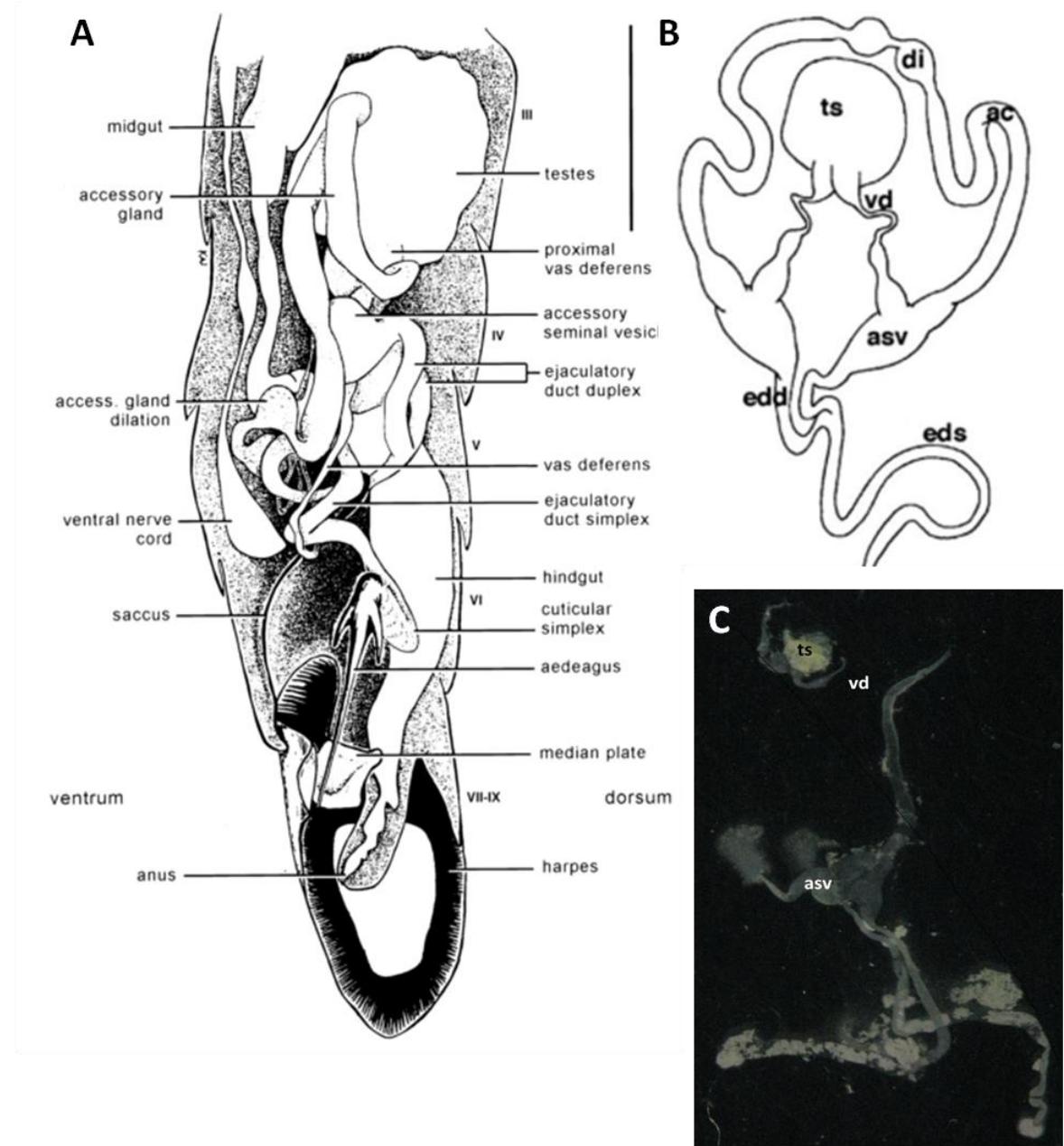


Figure 40. Reproductive structures of male *Plutella xylostella* (from Justus and Mitchell, 1999). A) Left lateral cuticle is excluded between ventral and dorsal midline to expose structures *in situ*. Roman numerals denote abdominal segments; bar = 500 µm. Testes are described as “fused into a sphere ca. 337 µm diameter and are situated dorsally in abdominal segments three and four.” B) A line drawing of the adult male reproductive system, black bar = 200 µm, after Justus and Mitchell, 1999. C) Light micrograph of the male reproductive system showing similarity with the line drawing (magnification ×10), no scale provided. During dissection the fine tubes are prone to tearing. Abbreviations: ac: accessory gland, asv: accessory seminal vesicle, di: dilation of accessory gland, edd: ejaculatory duct duplex, eds: ejaculatory duct simplex, ts: testes, vd: vas deferens. Not to the same scale. The testes are approximately 700 µm in diameter. The accessory glands are of particular interest and in contrast to other male Lepidoptera; they are joined distally.

Inevitably diagrams represent idealised anatomy; a photographic example of testes dissection is included (Figure 40B and C). Interestingly the testes are fused into a small sphere about 700 µm in diameter. Dissecting at different life-stages shrinkage and atrophy of the testes was observed, from pupa to 2-day-old adult.

4.3.3.4.1 OX4703 TESTES, RED OR NOT?

OX4703 2-day-old adult male testes were dissected (Methods 8.1.6) and screened for DsRed2 expression. The procedure was conducted on five males from each of the four lines plus an equal number of wild-type counterparts. Negligible auto-fluorescence was detected on the connective tissue surrounding the testis in both wild-type and transgenic. No obvious red fluorescence was observed.

Results of previous work indicates that there is relatively little variation in expression levels in testes of transgenes inserted into random genomic (autosomal) positions, so while chromatin domain organisation may facilitate testis-specific expression, it is not considered critical for ensuring normal gene expression levels (White-Cooper, 2010).

4.3.3.5 IMPROVED MAGNIFICATION TESTES FLUORESCENCE SCREENING

An alternative explanation is that the equipment used was not sensitive enough to detect low-level expression of the fluorescent proteins. Superior equipment was used on dissected testes looking at sperm bundles specifically in diamondback moth 2-day-old adults with increased magnification and fluorescence intensity (Olympus SZH10, Olympus UK Ltd.).

The duplex, vasa deferentia, seminal vesicles, and testes were exploded by dissection in PBS because there is a daily rhythm in the descent of mature sperm bundles from the testes to the seminal vesicles and then to the duplex for lepidopteran species. This ensured the sperm bundles were located, as often dissection sessions would last several hours.

The sperm bundle affects the light passing through in a distinctive way, with an appearance like the leaves of a book. However, for the first batch of dissections, Haust's solution was used to stain the nuclei, to facilitate eupyrene sperm identification (Figure 41).

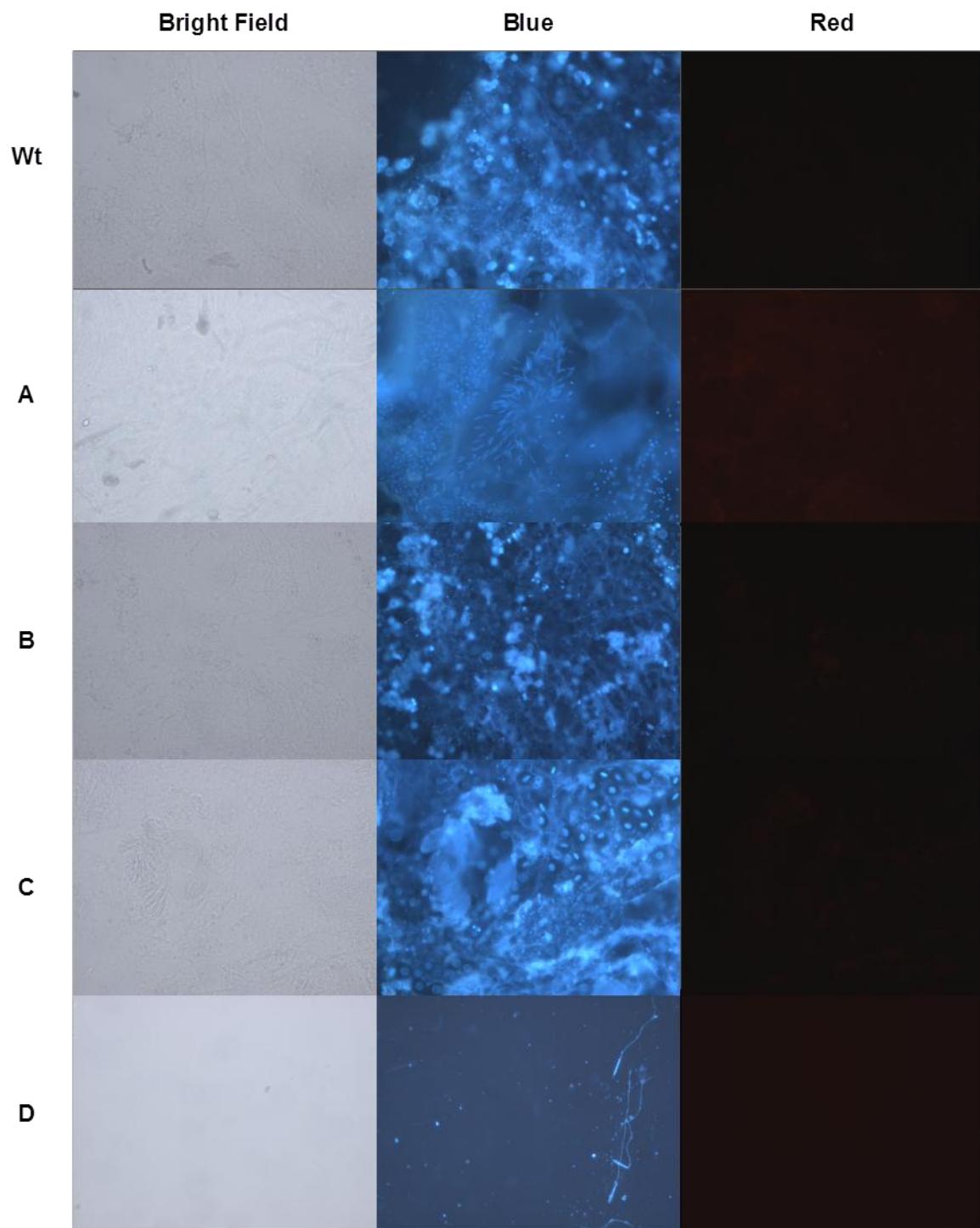


Figure 41. Diamondback moth testes dissected and exploded for inspection of fluorescence in sperm bundles. Haust's solution was used to dye the nuclei and pinpoint location of sperm bundles under blue filter. Fluorescent modes were then switched to red filter to detect DsRed2 expression in the sperm under the control of $\beta 2$ -tubulin promoter. The wild-type control is shown for comparison, wt. Transgenic lines A-C are shown with testes explosion, line D is shown with individual sperm visible at a higher magnification. The bright field image is included to show the necessity of Haust staining in order to locate sperm. No bleed through was seen nor any DsRed-sperm.

No fluorescent sperm were found in any of the transgenic lines suggesting very low expression levels of DsRed2, or none at all. The lack of detectable protein pointed to a transcription problem.

4.3.3.6 TRANSCRIPTION OF ENDOGENOUS *B2-TUBULIN* IN WILD-TYPE DIAMONDBACK

MOTH

4.3.3.6.1 DIAMONDBACK MOTH IN TWO-DAY OLD ADULTS

The lack of fluorescence could be explained by the promoter not switching on in a testes-specific and appropriately timed manner. Therefore RT-PCR was used to confirm that the *β2-tubulin* gene was being transcribed in wild-type male testes (Figure 42), as was expected. Testes were dissected out from 2-day-old wild-type diamondback moth adults.

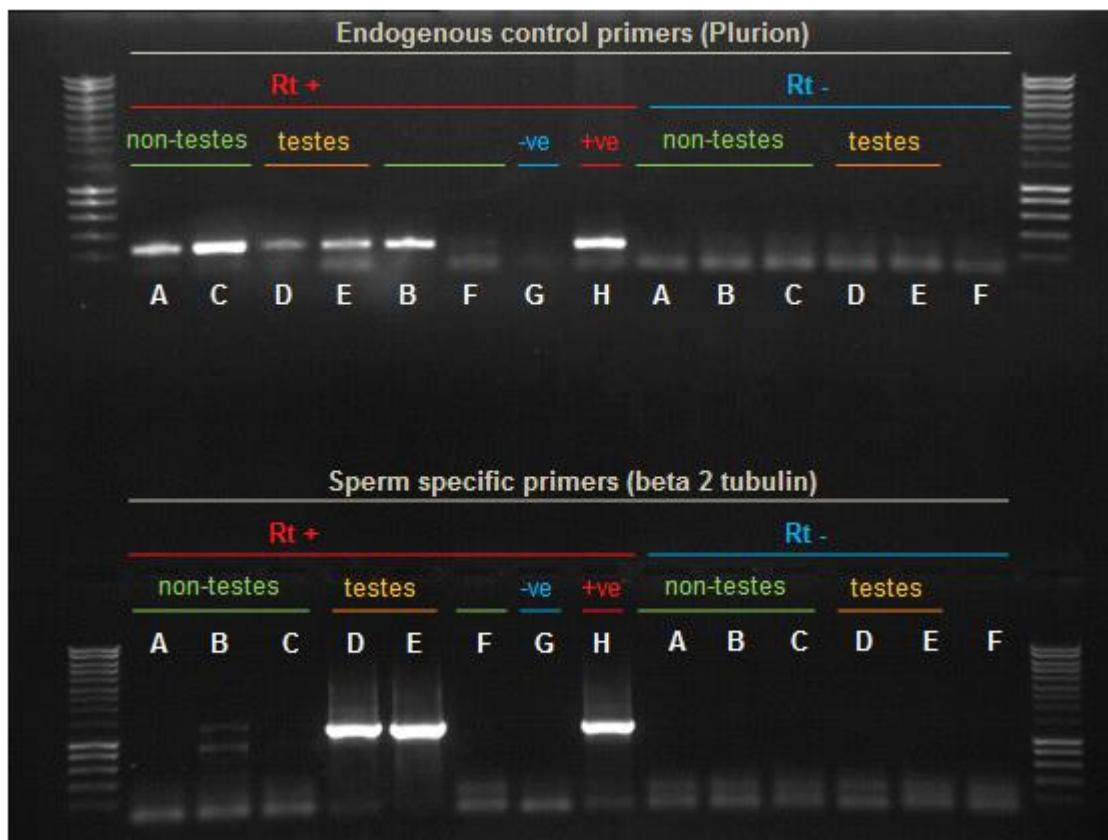


Figure 42. RT-PCR to detect putative $\beta 2$ -tubulin expression in different tissues dissected from wild-type diamondback 2-day-eclosed adult moths; specifically: A) female abdominal segment III, B) female carcass abdominal segments VI-IX, C) male carcass, abdominal segments VI-IX, D) male abdominal segment III, E) male testes, F) male non-testes major organ from abdominal segment III, G) negative control of water, H) positive control of testes derived cDNA from previous work. A control of PCR without the reverse transcriptase was also conducted. Tissue samples were designated testes containing or non-testes containing (control, both male and female included). The top half includes PCR products using control primers for endogenously expressed gene Plurion (vesicular acetylcholine transporter, VACHT) (GF#2570, Plurion F & GF#2571, Plurion R). The bottom half the $\beta 2$ -tubulin primers (or what we think are $\beta 2$ -tubulin primers) for diamondback moth (GF#2841, $\beta 2$ -tubulin across intron & GF#2753, $\beta 2$ -tubulin R). A positive control of cDNA from diamondback moth testis were also included (this positive control could be considered as a replicate diamondback moth tissue known to express $\beta 2$ -tubulin). Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The testes and third abdominal segment of the male (segment that contains the testes) were shown to contain transcripts of putative $\beta 2$ -tubulin. Those tissue samples that did not contain testes were negative for the $\beta 2$ -tubulin transcript. This result indicates that expression of this putative $\beta 2$ -tubulin is testis-specific.

4.3.3.6.2 DIAMONDBACK MOTH FOURTH INSTAR LARVAE

The lack of fluorescence could be explained by the promoter not switching on at a sufficiently early stage in testes. Therefore $\beta 2$ -tubulin RT-PCR was applied to fourth-instar larvae (Figure 43).



Figure 43. RT-PCR to detect putative $\beta 2$ -tubulin expression in different tissues dissected from wild-type fourth-instar larvae; specifically: A) female whole carcass, B) male non-testes containing carcass, C) male abdominal segment containing testes, D) testes, -ve) negative control of water. Tissue samples were designated testes containing or non-testes containing (control, both male and female included). The top half includes PCR products using control primers for endogenously expressed gene Plurion or VachT (Vesicular acetylcholine receptor transporter) (GF#2570, Plurion F & GF#2571, Plurion R; See Appendix 6). The bottom half the $\beta 2$ -tubulin primers (or what we think are $\beta 2$ -tubulin primers) for diamondback moth (GF#2841, $\beta 2$ -tubulin across intron & GF#2753, $\beta 2$ -tubulin R; See Appendix 6). The sperm-specific primers were included in the reaction tubes for C & D of the endogenous control primer reaction also by mistake. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

Testes-positive samples contained $\beta 2$ -tubulin mRNA transcripts (C & D in Figure 43). The amplicon (transcript by proxy) was not detected in non-testes containing tissues (A & B in Figure 43).

These two RT-PCR experiments suggest that *β2-tubulin* transcription coincides with spermatogenesis, which begins as early as the second instar in some lepidopteran species (Swallow and Wilkinson, 2002).

4.3.3.7 CONSTRUCT SEQUENCE INTEGRITY IN TRANSGENIC LINE

A possible explanation for the OX4703 not displaying the desired phenotype was that the line had been contaminated. OX4673 co-habited the lab at the same time and had the same Hr5-*ie1*-AmCyan marker. Genomic DNA was extracted from OX4703A, OX4703B and wild-type using PCR to investigate presence of construct specific DNA sequences (diluted OX4673 miniprep) (Figure 44).

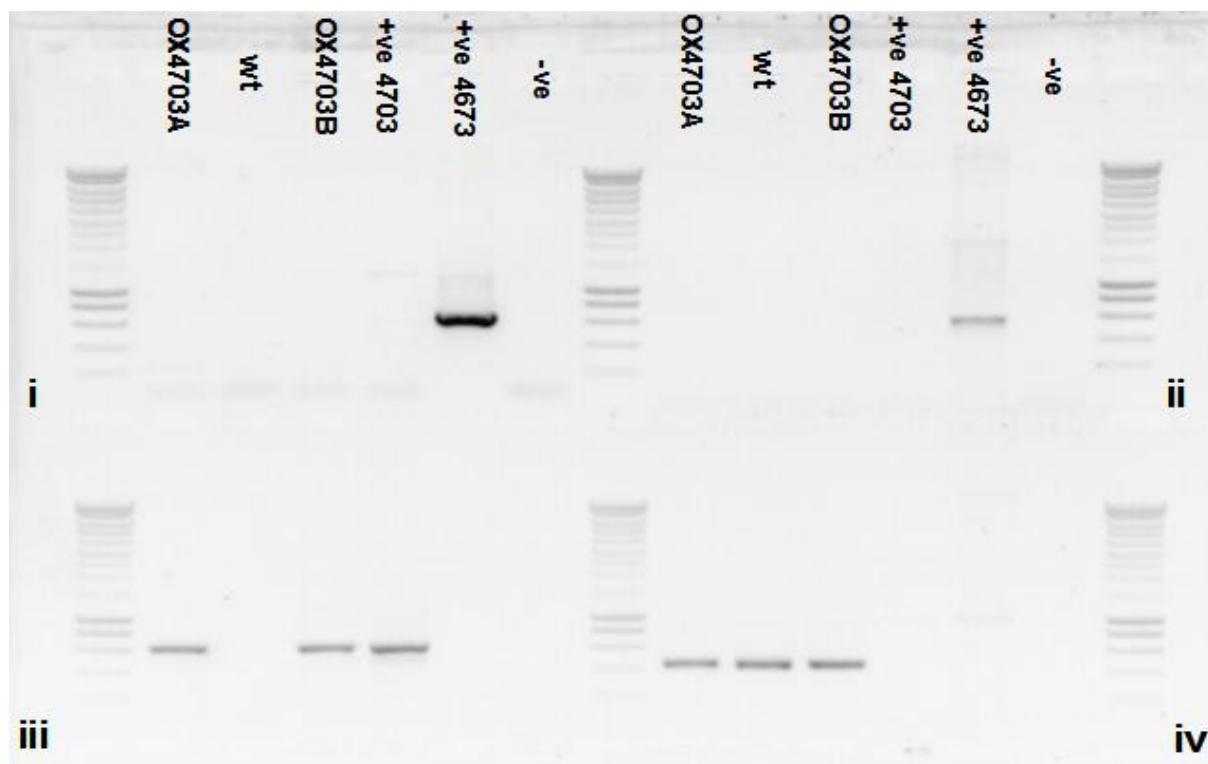


Figure 44. No contamination of OX4703 line with other constructs; the OX4703 insertion is real. A 1.2 % agarose gel (1.5 µl of ethidium bromide) run at 120 V for 30 minutes with the PCR products from four reactions (run as standard (4.6.10)) with unique primer pairs: i. OX4673 specific – primers TD2797 and TD1830 expected 648 bp., ii. OX4673 specific – primers TD1884 and TD1884 expected 655 bp., iii. OX4703 specific – primers TD3275 and TD1406 expected 606 bp., iv. Endogenous housekeeping gene – primers SS1487 and SS1488 expected 458 bp. The band presence on the positive control of the reactions i) and ii) demonstrated the primers functionality. No other bands suggested there has been no contamination of the OX4703A and B lines from OX4673 by mating since the OX4703 lines were established. The primers specific for OX4703 amplified DNA in reaction iii) for both samples OX4703A and B. The amplification of the endogenous gene suggested the integrity of the extracted gDNA. Smart ladder included on the flanks for PCR product length estimation. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

No contamination was evident.

The amplicon for OX4703A & B was sequenced and compared with the designed construct sequence (Methods), revealing that the sequence was as expected with a two nucleotide change, possibly due to PCR / sequencing error.

Although the fluorescent protein may not be expressed at detectable levels this does not rule out the absence of protein; therefore a more sensitive molecular method was required to detect the precursor to protein. Male (transgenic and wild-type) and female carcasses (transgenic) were tested for the presence of detectable levels of mRNA coding for the β 2-tubulin regulated

genes. For OX4703 we used primers specific for the *Dmprot* and DsRed2, unique to the construct. For each sample four pupae were sexed and stored at -80°C. Samples were analysed for presence of transgene and $\beta 2$ -tubulin transcripts by RT-PCR (Figure 45).

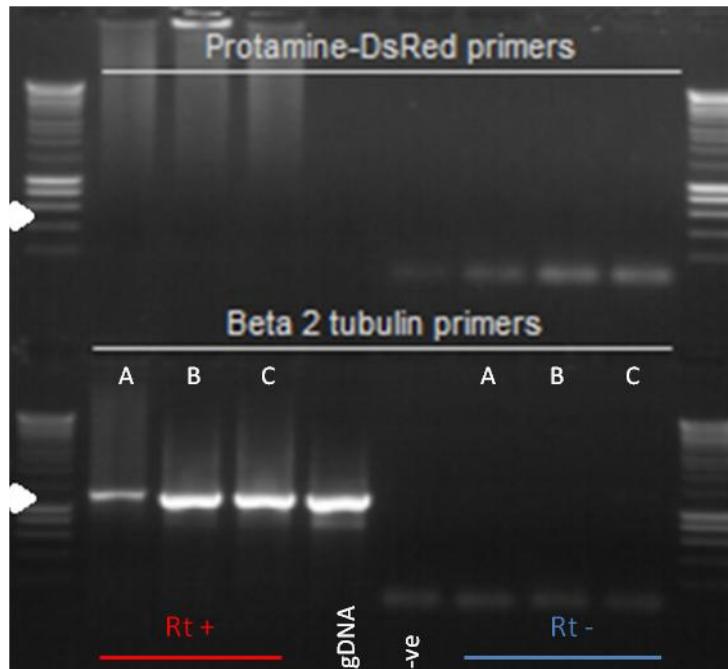


Figure 45. Determining the presence of the transgene transcript in OX4703 when compared to the endogenous $\beta 2$ -tubulin transcript. The columns are constant in terms of the nucleic acid sample type but each row has different primer pairs. The samples correspond to female transgenic carcass cDNA (A), male transgenic carcass cDNA (B) and wild-type male carcass (C). The genomic DNA was extracted from OX4703A (gDNA). The negative (-ve) control represents a reaction with no nucleic acids. The three left-most columns are from the RT-PCR of the samples RNA extracts with the reverse transcriptase step creating cDNA, the three right-most columns represent the controls without the reverse transcriptase step leaving RNA. The white arrows represent expected amplicon size given the primer pair used in the set of reactions. Primers for protamine and DsRed2 were available for use in the PCR of the cDNA obtained from the RT-PCR. A control of endogenous genes was used concurrently including the endogenous $\beta 2$ -tubulin gene to control for the endogenous $\beta 2$ -tubulin promoter working. If protamine/DsRed2 mRNA were present then the appropriate bands would be detected on the agarose gel in the transgenic male (sample B) but not the transgenic female (sample A) nor the wild-type male (sample C). A positive control of genomic DNA extracted from the OX4703A was used to control for primers working. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The endogenous $\beta 2$ -tubulin was transcribed in both males as well as being present in the genomic DNA of the transgenic insect. The female amplicon was not expected but is larger than the male suggesting a different tubulin isoform was amplified. This evidence suggests that the endogenous $\beta 2$ -tubulin gene has been transcribed. In hindsight, primers including introns are

preferred but we can rule out DNA contamination with the reverse transcriptase negative control.

Assuming a functional $\beta 2$ -*tubulin* promoter, one would expect transcription of the exogenous protamine-DsRed2 gene. No transgene amplicons were detected. These primers had been shown to be compatible in RT-PCR targeting *Dmprot-DsRed2* in Medfly transgenic insects. Independently this experiment is inconclusive but in combination with previous data suggests the transgene is not being transcribed.

4.3.3.8 REASONS FOR OX4703 FAILURE - SIMULATION OF THE CONSTRUCT USING TINKERCELL

To aid design and troubleshooting, a range of computer aided design (CAD) approaches can be used (Chandran *et al.*, 2009). The model aids visualisation and explanation of the sperm-specific module failure. Parameter values for molecule kinetics associated with the construct can also be applied, if for example there were a desire to optimise the promoter region for the desired phenotypic effects. Figure 46 describes the gene network and the regulatory elements involved.

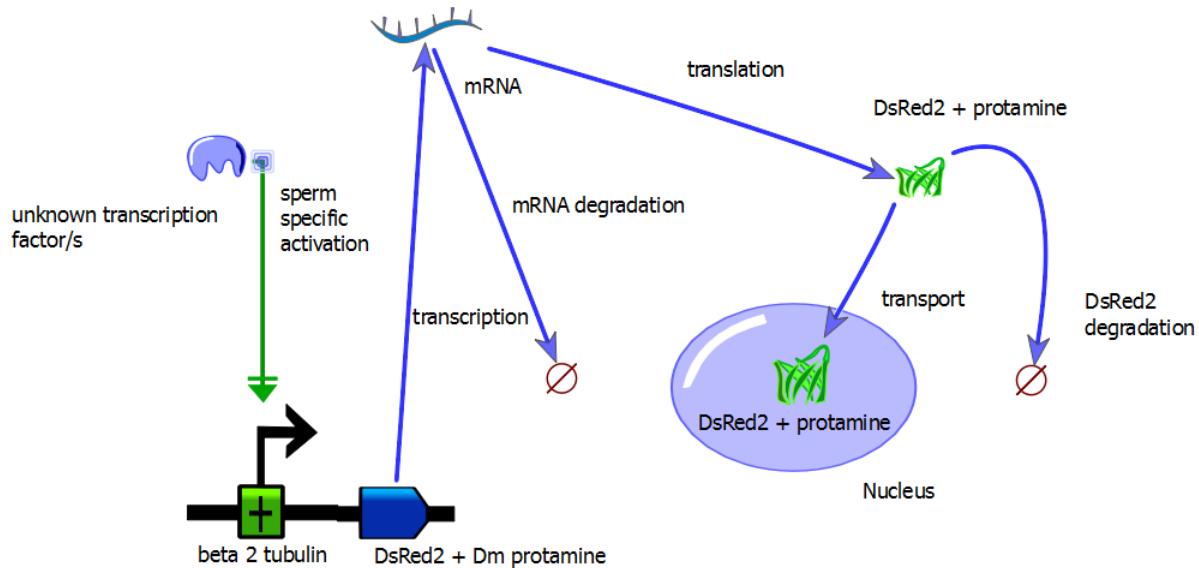


Figure 46. The genetic circuitry associated with the sperm specific transcription and expression of a reporter gene to then be transported into the nucleus. This represents the first step towards a paternal effect system. The pathway nature of the network suggests that a fault with any of the nodes (symbols) or at any of the vertices (arrows) would prevent the desired phenotype. Changes in relative quantities of the molecules over time can be simulated with differential equations and the Hill function (SysMIC course, www.sysmic.ac.uk, enrolled 2013, completed 2015).

Qualitative values can be placed for the initial conditions [i.e. the transcription factors that provide sperm specific transcription will be present in males but not in females, so is assigned a value of 1 (from a functional perspective)] of the variables and approximate functions applied to determine the effects of any faulty stages of the system (Figure 47). Combining the simulated results with the experimental data provides an insight as to how the system is failing and a basis for designing a system that is more likely to work.

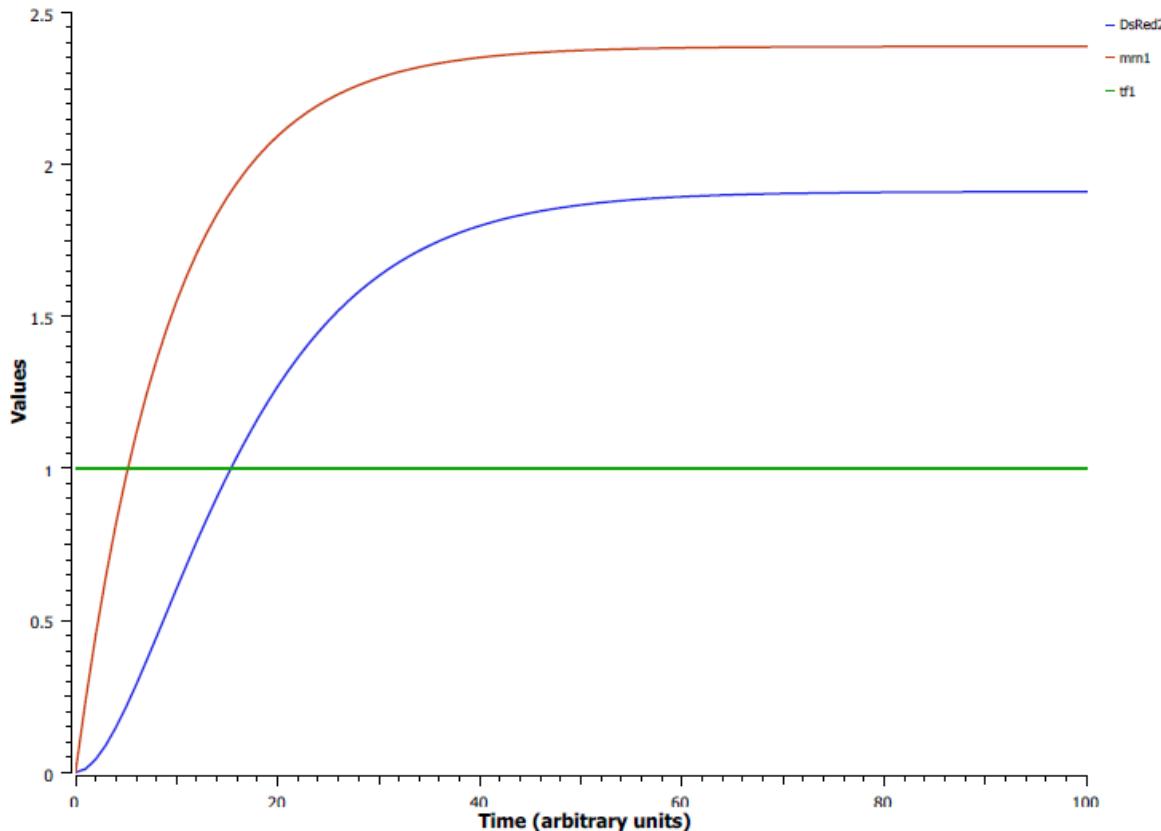


Figure 47. The transcription factor (tf1) is assumed constant (expressed simultaneously and instantly at maximum in sperm cells at the late larval life stage) shown by the green line. Tf1 binds to the $\beta 2$ -tubulin promoter inducing transcription of mRNA (mrn1 – the red line). The mRNA is translated into the reporter protein (DsRed2 - blue line). The maximum occurs when rate of production equals the rate of degradation.

Changing the parameters of the model elucidated possible reasons for the desired phenotype not occurring:

- Putative promoter region does not promote transcription initiation.
- 5' UTR of the $\beta 2$ -tubulin and protamine interferes with transcription elongation.
- mRNA degrades faster than it can be translated.
- mRNA translation fails.
- DsRed2 and protamine fusion non-functional or degrades.
- DsRed2 and protamine fusion not transported.
- Splicing failure of Dm intron causing DsRed2 to be out of frame.

The unknown transcription factors are present in the male sperm cells at the fourth instar larval and adult life stages (evidenced by the transcription profile of the $\beta 2$ -*tubulin* gene in wild-type diamondback moth). The putative $\beta 2$ -*tubulin* promoter is activated and promotes the transcription of the DsRed2 and protamine fusion protein. As no mRNA for this fusion protein were detected then it can be inferred that the pathway stops at or before the transcription stage.

4.3.4 OX4804 – TESTING 5' UTR *DMPROT* INTERFERENCE?

4.3.4.1 TROUBLESHOOTING OX4703

A new construct, OX4804, was designed after discussion of the likely cause/s of OX4703 failure. A problem with the transcription machinery may have been caused by the 5' UTR hybrid of the protamine and the $\beta 2$ -*tubulin* promoter conflicting ('biological crosstalk'). The 5' UTR of the *Dmprot*, the transcription start and the intron were removed; the protamine functionality is desired but none of the regulatory affects. Intron removal was recommended by Helen White-Cooper (personal communication), as it was suggested that experience has shown intron inclusion in testis-specific genes impairs transcription, opposite to her experience with somatic transcription. The expression of *Dmprot*-DsRed2 is regulated by the diamondback moth $\beta 2$ -*tubulin* promoter in this construct with the 5' UTR and transcription start included (see Appendix 9.1.4 for design and synthesis details).

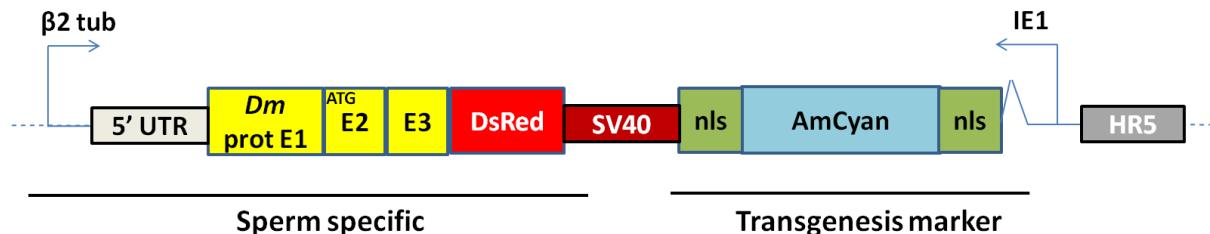


Figure 48. A simplified representation of the genetic construct OX4804, a more detailed description combined with component summary can be found in appendix 9.1. The β 2-tubulin promoter was derived from diamondback moth genome inspection and successfully cloned. The main transgenesis marker is provided by the immediate early promoter (Hr5-*ie1*) with AmCyan-SV40. The sperm-specific components were regulated by about 1 kb upstream of diamondback moth β 2-tubulin. The protamine, this time minus the 5' UTR, was derived from *Drosophila melanogaster* (yellow) and was fused to DsRed2-SV40. The protamine consists of exon 2, an intron and exon 3. The dashed line represents other components of the circular construct. Throughout this thesis genes are typically tailed by SV40 3' UTR unless stated otherwise, the SV40 is not bi-directional but only one copy is shown here for brevity.

4.3.4.2 TRANSGENESIS OF OX4804 IN DIAMONDBACK MOTH

A total of 1212 embryos were injected with construct OX4804. The 388 injection survivors were set up in crosses of approximately 40 G₀ and 10 wild-type females. G₁ insects were screened for the Cyan marker. The promoter-marker phenotype (Hr5-*ie1*-AmCyan) was expected to be similar to OX4703.

Table 12. The injection data summary for the construct OX4804 in diamondback moth. Survival was 30%, just over the median suggesting an improvement relative to OX4703-injected survivors. Proportion survival and transformation efficiency given in brackets.

| Species | Construct | Embryo injections (n ₁) | Injection survivors (G ₀) | Independent transgenic lines (X) |
|----------------------------|-----------|-------------------------------------|---------------------------------------|----------------------------------|
| <i>Plutella xylostella</i> | 4804 | 1212 | 388 (0.55) | >3 (<0.01) |

Fluorescent marker in transformants was similar to that in OX4703-transformed lines (Figure 49). Again it was noted how screening the pupae dorsally was preferred to avoid false negatives.

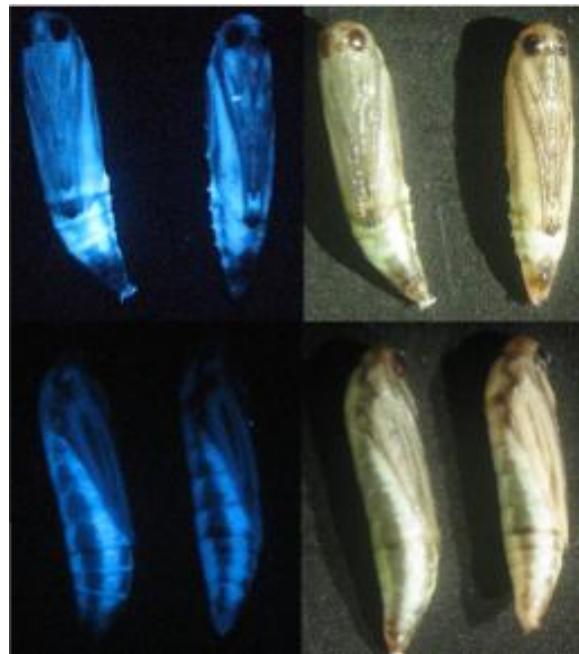


Figure 49. Fluorescence microscopy image of a putative OX4804 transformant. The top row shows two pupae viewed ventrally and the bottom row is the same two pupae viewed dorsally. The OX4804 construct uses the AmCyan marker; the left column is viewed through the fluorescence microscope using Cyan filters and the right column is under white light conditions. The pupae are approximately 6 mm in length. Detection of transgenics depends on inspection of the dorsal side adjacent to the spiracles where higher concentrations of fluorescent protein can be seen.

4.3.4.2.1 CHECKING FOR INSERTION NUMBER USING WILD-TYPE CROSSES

Eighteen unique crosses were set up, each with three wild-type mates (male transgenics were used where possible). Observed ratios of fluorescent and non-fluorescent offspring were screened for at G₂, and compared with expected ratio given Mendelian inheritance (Table 13).

Table 13. Putative transgenic G₁ males were separated and labelled as different putative lines. One transgenic male was crossed with three wild-type females. Eggs were collected and reared to pupae. The frequency of transgenic to non-transgenic pupae were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios [or a proportion significantly greater than 50% transgenic (indicated by AmCyan fluorescent protein marker) to non-transgenic phenotype]. Evidence for multiple insertions of the transgene in any line was weak. Several crosses yielded more non-transgenic offspring than expected. Proportions given in brackets.

| Line | Transgenic | Non-transgenic | χ^2 (2.d.p) | p (2.d.p) |
|-------------|-------------------|-----------------------|------------------------------------|------------------|
| A | 27 (0.79) | 34 (0.21) | 0.80 | 0.37 |
| B | 24 (0.47) | 27 (0.53) | 0.18 | 0.67 |
| C2 | 2 (0.12) | 15 (0.88) | 9.94 | 0.00 |
| C3 | 32(0.36) | 58 (0.64) | 7.51 | 0.01 |
| C4 | 35 (0.64) | 20 (0.36) | 4.09 | 0.04 |
| C5 | 28 (0.41) | 40 (0.59) | 2.12 | 0.15 |
| D1 | 83 (0.43) | 117 (0.57) | 5.78 | 0.02 |
| E1 | 42 (0.30) | 98 (0.70) | 22.40 | ~0 |
| G | 35 (0.48) | 38 (0.52) | 0.12 | 0.73 |
| H | 4 (0.25) | 12 (0.75) | 4.00 | 0.05 |
| I | 3 (0.05) | 58 (0.95) | 49.59 | ~0 |

The persistence of the fluorescent phenotype in expected Mendelian ratios supported the notion of a single insertion (some crosses showed significantly non-Mendelian ratios due to a greater frequency of non-transgenic offspring).

4.3.4.2.2 CONFIRMING CONSTRUCT INHERITANCE USING GDNA EXTRACTION AND PCR
WITH OX4804 SPECIFIC PRIMERS

The genomic DNA of two putative OX4804-transformed lines were extracted and run in a PCR with primers specific to the construct (Figure 50).

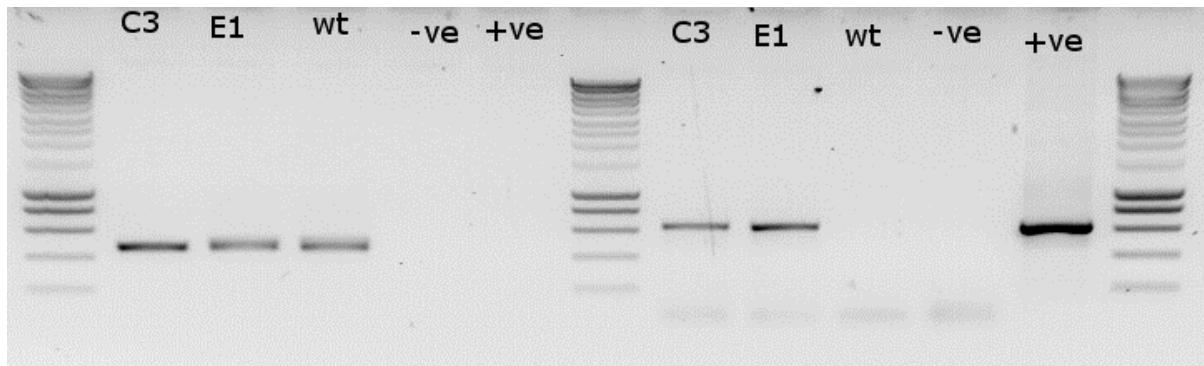


Figure 50. Molecular confirmation of inheritance of the construct OX4804. A 1.2 % agarose gel (1.5 μ l of ethidium bromide) run at 120 V for 30 minutes with the PCR products from two reactions with unique primer pairs: left; Endogenous housekeeping gene *doublesex* – primers SS1487 and SS1488 expected 458 bp; OX4804 specific – primers TD3275 and TD1406 expected 606 bp. Smart ladder included on the flanks for PCR product length estimation. The gDNA samples were extracted from putative transgenic lines OX4804C3 (C3) and OX4804E1 (E1) and wild-type (wt). The negative control of water instead of DNA (-ve). The positive control of OX4804 diluted injection mix (+ve) were also included. Colour was inverted using Paint. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

4.3.4.3 TESTES DISSECTION AND FLUORESCENCE MICROSCOPY

4.3.4.3.1 ADULTS

Wild-type was dissected as a negative control for comparison. Fluorescence was assessed at the whole testes level (Figure 51) and for sperm bundles (Figure 52).

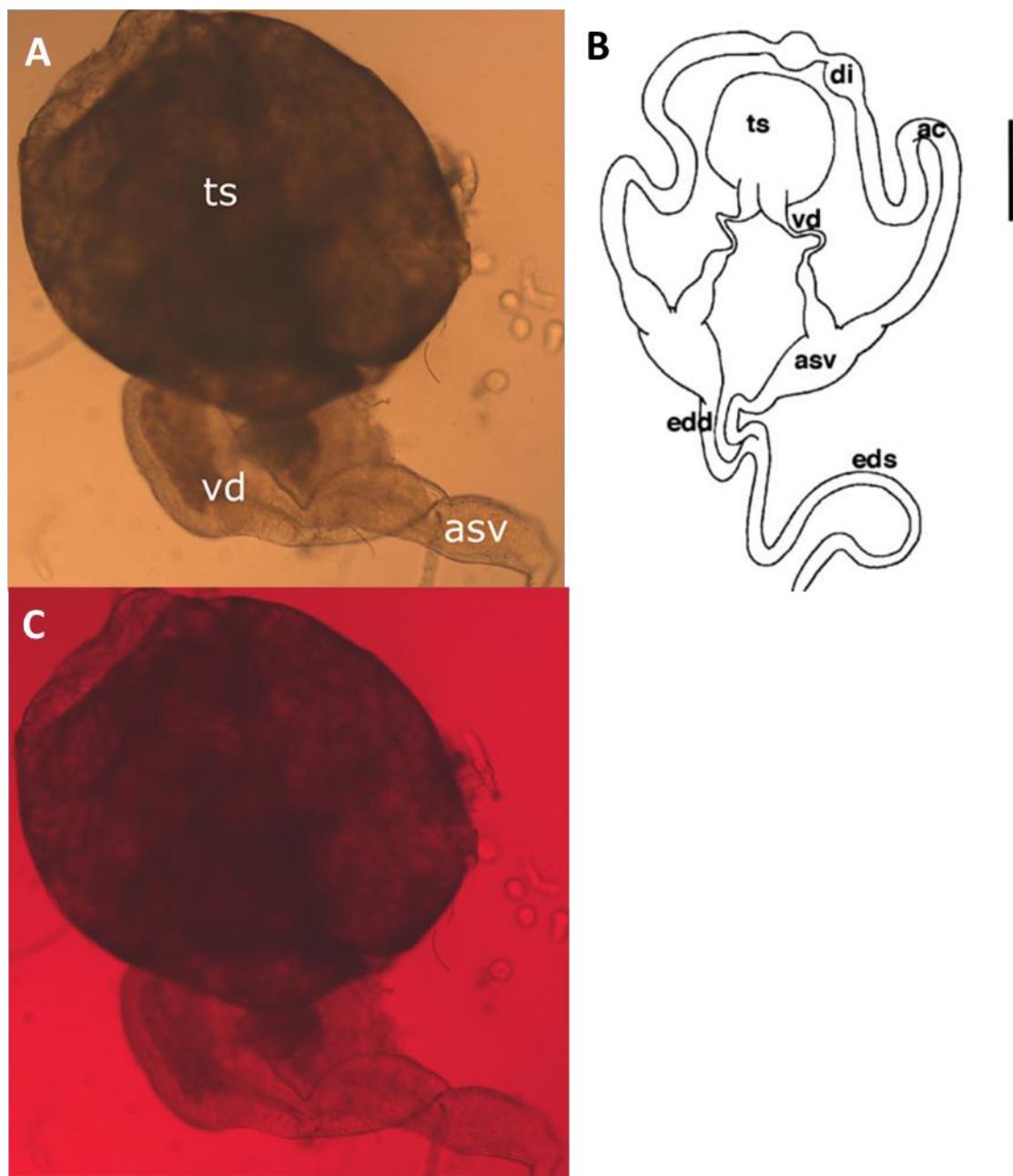


Figure 51. (A) Light micrograph of the male reproductive system showing similarity with the line drawing (Magnification $\times 40$), no scale provided, although the diameter of the testes were approximately 300 μm for reference. During dissection the fine tubes are prone to tearing; shearing the testes and the vas deferens from the rest of the male reproductive system, separating at the narrowest point. The Image was taken using DIC phase contrast with the sample in testes buffer on a glass slide under a glass cover slip. (B) A line drawing of the male reproductive system, black bar = 200 μm , taken from Justus and Mitchell, 1999. Abbreviations: ac: accessory gland, asv: accessory seminal vesicle, di: dilation of accessory gland, edd: ejaculatory duct duplex, eds: ejaculatory duct simplex, ts: testes, vd: vas deferens. Not to the same scale. (C) Identical sample from A with incident excitation light and filters specific for DsRed2, for annotation see A.

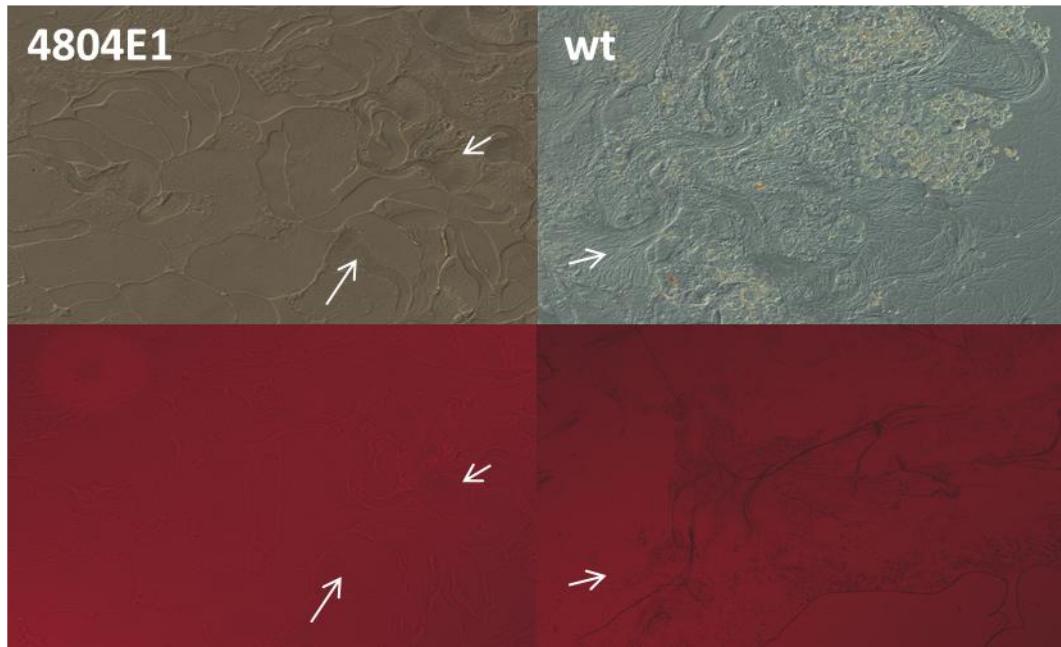


Figure 52. The left pictures are the same sample of OX4804E1 testes taken with phase contrast DIC (top) and appropriate incident light and DsRed2 filter. The right pictures are for wild-type as a control to allow comparison. No fluorescence is detectable in either sample. The white arrows show sperm bundles, for comparison between visualisation methods (rows).

Ten transgenic lines were screened (OX4804B, A, E2, C4, I1, E1, C3, D1, G1 and C5) with at least two males from each line dissected and inspected for DsRed2 at detectable levels in the testes or sperm bundles. No fluorescent testes or sperm bundles were detected.

4.3.4.3.2 FOURTH INSTAR LARVAE

$\beta 2$ -tubulin expression was shown to occur in the fourth-instar larval and 2-day-old adult stages (Figure 42 & Figure 43). Accordingly, dissection of testes and screening for DsRed2 from these life stages were conducted. Figure 53 shows the expected cross-sectional anatomy and different cell type positions in the testes. A development gradient is seen from right to left, less to more developed, towards maturity of the sperm.

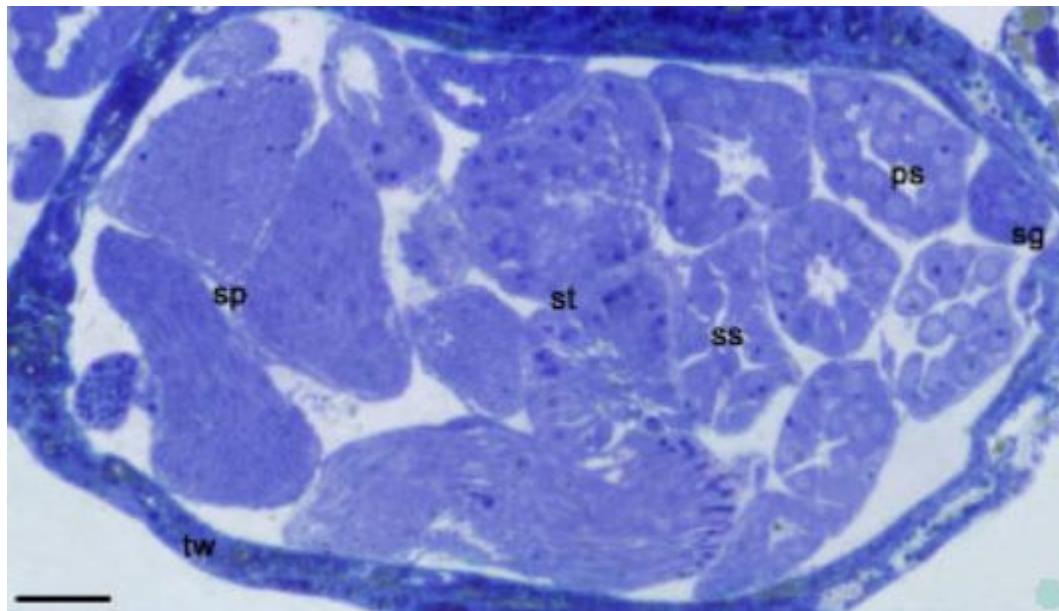


Figure 53. Photomicrograph of testes morphology in *Plutella xylostella* larva. Tw: testicular wall; sg: spermatogonia; ps: primary spermatocytes; ss: secondary spermatocytes; st: spermatids; sp: sperm. Scale bar: 50 µm (Tanaka & Tagashira, 1998).

The fourth instar larvae testes were examined at 100x magnification for DsRed2. Five lines were chosen at random with up to seven insects from each line dissected and examined for fluorescent testes. Testes were dissected and exploded prior to viewing.

Table 14. Fourth instar larvae males of five OX4804-transformed lines had testes removed and examined for DsRed2. No fluorescence detected. The number of testes examined varied between lines due to escapees and loss of testes during dissection.

| OX4804 Line | Pairs of testes examined | Fluorescence detected |
|-------------|--------------------------|-----------------------|
| A | 3 | 0 |
| C3 | 3 | 0 |
| E1 | 7 | 0 |
| G | 2 | 0 |
| H | 2 | 0 |

No DsRed2 was detected in any of the transgenic lines when compared to the wild-type control. The mRNA of these lines was examined to determine if β 2tubulin-Dmprot-DsRed2 transcription occurred.

 4.3.4.4 *B2-TUBULIN*-REGULATED DSRED2 AND PROTAMINE TRANSCRIPTION IN
 OX4804 TRANSFORMANTS

Molecular detection of mRNA transcripts is likely to be more sensitive than visual detection of fluorescent proteins, if transcription and expression were low. Accordingly RNA was extracted from three different lines of OX4804 and assessed for transgene specific transcript using RT-PCR (Figure 54).

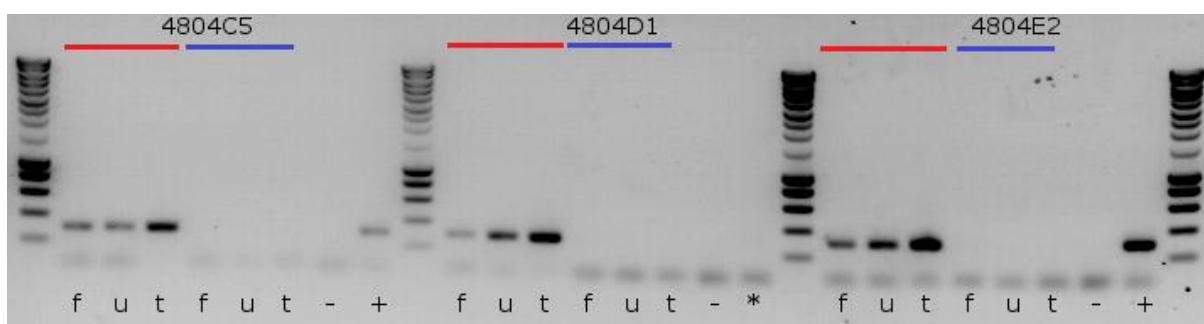


Figure 54. Assessing testes specific transcription of the transgene in OX4804. RT-PCR products of cDNA with primers specific for the OX4804 transgene, DsRed2 and *Dm* protamine (primers TD825 and TD3275). The cDNA were derived by RT-PCR of three different types of insect samples (male fourth instar larvae: t – abdominal segments including testes, u – from the same insect the upper half of the insect remaining after bisection above the testes; female late larvae, f – the entire insect) of three different transgenic lines (not necessarily independent insertions: OX4804C5, OX4804D1, OX4804E2). The PCR products were run on a 1.2% agarose TEA gel at 120V for 30 minutes. The red lines show the samples from RT-PCR and the blue from no reverse transcriptase RT-PCR. The * represents an anomalous result; prior to loading in the thermocycler it was noted that the well containing the positive control for 4804D1 had ice in it and reduced volume compared to the other wells. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1 kbp, then by 500 bp to 3 kbp (Eurogentec, SF Smartladder).

Given the total RNA concentration from which the cDNA were derived the near saturation of the PCR for the testes, (at 35 cycles, the band appears thick and bright relative to the other tissues but not to the ladder) is indicative of higher rates of transcription in these tissues relative to the female and the upper half of the male, which were somewhat equivalent in cDNA copy number (as a proxy for transcript copy number). Only one male per line was included so apparent inter-strain variation could be caused by inter-insect or intra-sample variation. The amplicon concentration is equivalent to that of the 600-bp rung at 60 ng per band.

4.3.4.5 OX4804 CRITIQUE

This information, combined with the fluorescence microscopy inspection of the testes of OX4804 transformants, suggests that although some transcript is being produced not enough is being transcribed and / or translated for DsRed2 detection. This is a concern as a heterodimeric nuclease under control of the current promoter may be inadequately transcribed for the desired paternal effect phenotype. The leaky transcription in non-testes tissues suggests we were missing an endogenous non-coding DNA repressor region.

Compared with other examples in the literature (Catteruccia *et al.*, 2005; Scolari *et al.*, 2008; Zimowska *et al.*, 2009), the $\beta 2$ -tubulin promoter normally provides strong testis-specific transcription and subsequent translation of associated fluorescent proteins. This suggests that perhaps the 955-bp putative promoter region included in OX4804 (ignoring the 5' UTR) does not contain all the necessary upstream elements of the $\beta 2$ -tubulin gene required for the desired transcription efficiency. A new construct was designed to include a larger fragment of the upstream region.

A further concern related to the fidelity of the diamondback moth sequence data that we were using to design our constructs. The shotgun sequencing method can struggle with gene duplicates or repeated sequences (Franca *et al.*, 2002). I therefore checked the 5'UTR sequence of the $\beta 2$ -tubulin gene.

RNA was extracted from wild-type male abdominal segments containing the testes and the 5' UTR elucidated using Rapid Amplification of cDNA Ends (RACE) (Appendix 9.2). Comparison of the 5' RACE sequence with the Liverpool data through alignment in Vector NTI showed no significant differences (data not shown).

4.3.5 OX5067 – BIGGER IS BETTER?

4.3.5.1 TROUBLESHOOTING OX4804

It was decided to create a new construct with extended upstream sequence (~ 3 kbp) of the $\beta 2$ -*tubulin* gene included. The data suggested the promoter region or an enhancer was absent from the putative promoter region of OX4804 (1132 bp upstream of start ATG codon). It was hypothesised the addition of this region would increase transcription and translation producing detectable levels of DsRed2 in the sperm of a transgenic male.

The sequence data provided by Liverpool was aligned against the Chinese group's sequence data and was equivalent. No open reading frames of known genes were detected within 5 kb of the $\beta 2$ -*tubulin* gene. The new promoter region was designed to extend to a C/G-rich region to facilitate primer design to amplify further upstream of the $\beta 2$ -*tubulin* gene than on OX4804. The construct was designed and synthesised (Appendix 9.1.6).

The new promoter region and insertion brought the total putative promoter region to ~ 4.5 kbp in length. Other than this extension the new construct OX5067 was identical to OX4804 (Figure 55).

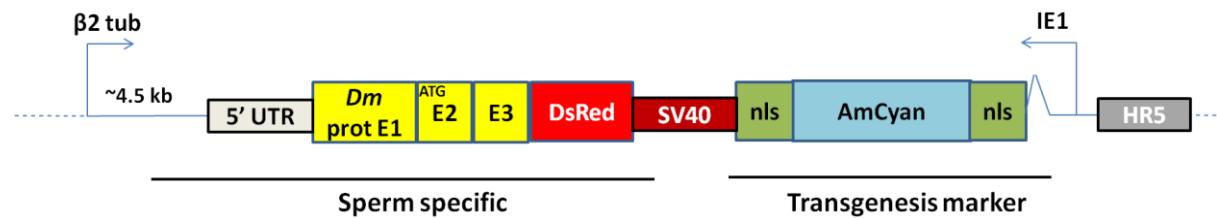


Figure 55. A simplified representation of the genetic construct OX5067, a more detailed description combined with component summary can be found in the Appendix. The $\beta 2$ -*tubulin* promoter was successfully extended relative to OX4804 and OX4703. The main transformation marker is provided by the immediate early promoter (Hr5-ie1) with AmCyan-SV40. The sperm specific components were regulated by about 4.5 kb upstream of diamondback moth $\beta 2$ tubulin. The protamine, this time minus the 5' UTR, was derived from *Drosophila melanogaster* (yellow) and was fused to DsRed2-SV40. The dashed line represents other components of the circular construct. Throughout this thesis genes are typically followed by SV40 3' UTR unless stated otherwise.

4.3.5.2 TRANSFORMATION WITH OX5067

A total of 1781 embryos were injected with construct. G₁ insects were screened for the AmCyan marker.

Table 15. The injection data summary for the construct OX5067 in diamondback moth. Survival was 31% - just over the median - suggesting an improvement relative to OX4703 survival. Proportion survival and transformation efficiency given in brackets.

| Species | Construct | Embryo injections (n1) | Injection survivors (G0) | Independent transgenic lines (X) |
|----------------------------|-----------|------------------------|--------------------------|----------------------------------|
| <i>Plutella xylostella</i> | 5067 | 1781 | 546 (0.31) | 6 (0.01) |

A total of 32 transgenic G₁ insects were found of the total 11,296 screened. To manage the workload some of the transgenic insects were discarded. The transgenesis module used was identical to that of OX4703 and OX4804 and similar phenotypes and phenotypic variation were observed in transgenic lines (Figure 64).



Figure 56. Fluorescence microscopy image of a putative transgenic OX5067 line designated G2. Three pupae are viewed dorsally with the wild-type (wt) on the left and two OX5067G2, of a similar phenotype under white light, on the right. The OX5067 construct uses the Hr5-*ie1*-AmCyan marker; the top row is viewed through the fluorescence microscope using Cyan filters and the bottom row is under white light conditions. The pupae are approximately 6 mm in length. Detection of transgenics depends on inspection of the dorsal side adjacent to the spiracles where higher concentrations of fluorescent protein can be seen. Variation within a line is shown here; however the marker is reliable when screened for by humans.

4.3.5.3 BACKCROSSES OF PUTATIVE LINES

A backcross was conducted on putative transgenic lines, and progeny screened for the transformation marker to assess for multiple insertions (Table 7). Deviation towards a greater ratio of transgenics would suggest more than one insertion of the construct into the ancestral G₀ gamete. The evidence from the crosses supported the null hypothesis of one insertion for all the lines except B1, which may have had multiple insertions. A fitness cost may contribute to the lower-than-expected number of transgenic individuals in the cross with line D1.

Table 16. Putative transgenic G1 males derived from unique G₀ crosses were separated and labelled as different putative lines. One transgenic male was crossed with five wild-type females (or one female with two wild-type males if no females were produced by a particular G₀ cross). Eggs were collected and reared to pupae. The frequency of transgenic to non-transgenic pupae were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios (or a proportion significantly greater than 50% transgenic (indicated by AmCyan fluorescent protein marker) to non-transgenic phenotype). Evidence for multiple insertions of the transgene in any line was weak except for B1. Cross D1 had offspring that were significantly more non-transgenic than expected.

| Line | Transgenic | Non-transgenic | X ² (2.d.p) | P (2.d.p) |
|------|------------|----------------|------------------------|-----------|
| A1 | 32 (0.52) | 30 (0.48) | 0.065 | 0.80 |
| A2 | 28 (0.47) | 30 (0.53) | 0.069 | 0.79 |
| A3 | 39 (0.49) | 41 (0.51) | 0.05 | 0.82 |
| B1 | 101 (0.66) | 52 (0.34) | 15.69 | < 0.001 |
| C1 | 75 (0.46) | 89 (0.54) | 1.20 | 0.27 |
| C3 | 54 (0.52) | 49 (0.48) | 0.24 | 0.62 |
| C4 | 14 (0.45) | 17 (0.55) | 0.29 | 0.59 |
| D1 | 27 (0.26) | 75 (0.74) | 22.59 | < 0.001 |
| E1 | 15 (0.36) | 27 (0.64) | 3.43 | 0.064 |
| F1 | 17 (0.43) | 13 (0.57) | 0.53 | 0.47 |

For insertion confirmation, gDNA was extracted and a PCR run using appropriate primers targeting the transgene. A randomly selected fourth-generation pupae (randomly selected from

the Petri dish in which the line was screened) from four random lines was chosen alongside a wild-type sample (Figure 57). The putative lines were confirmed as carrying DNA sequences present in OX5067.

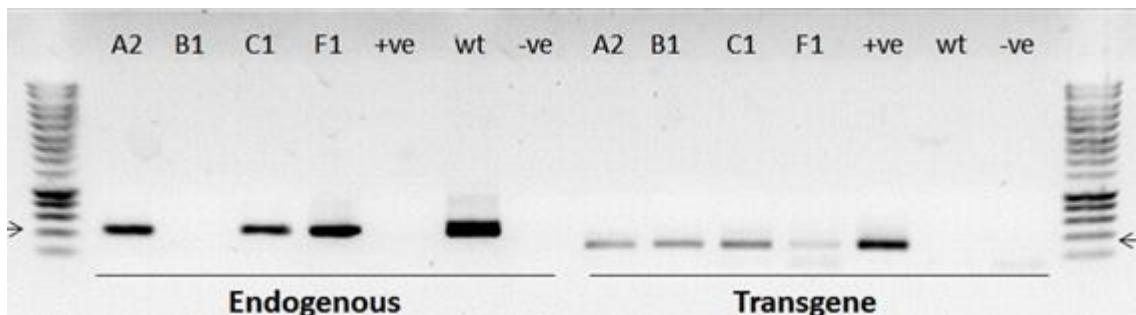


Figure 57. Molecular confirmation of the insertion of the construct OX5067 in putative lines. A 1.2 % agarose gel (1.5 µl of ethidium bromide in 100 ml) run at 120 V for 30 minutes with the PCR products from two reactions with unique primer pairs: left; Endogenous housekeeping gene *doublesex* – primers SS1487 and SS1488 expected 458 bp (left arrow on ladder); OX5067 transgene (across DsRed2 and Dmprotamine) – primers TD3275 and TD825 expected 291 bp (right arrow). Smart ladder included on the flanks for PCR product length estimation. The gdNA samples were extracted from pupae of the putative transgenic lines OX5067A2 (A2), OX5067B1 (B1), OX5067C1 (C1) and OX5067F1 (F1) and wild-type (wt). The negative control of water instead of DNA (-ve). The positive control of OX4804 diluted injection mix (+ve) were also included (the constructs only differed in the beta 2 tubulin promoter region). All lines were positive for the transgene associated sequence (DsRed2 and *Dm* protamine-specific region not found in wild-type gdNA or in the negative control of water instead of gdNA). Surrounding lanes provide ladder reference. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1 kbp, then by 500 bp to 3 kbp (Eurogentec, SF Smartladder).

4.3.5.4 TESTES DISSECTION AND FLUORESCENCE MICROSCOPY

4.3.5.4.1 LARVAE

Ten randomly selected third and fourth instar larvae were dissected and the testes and sperm bundles examined for fluorescence, none was detected.

4.3.5.4.2 TWO-DAY OLD ADULTS

Dissection and examination of sperm bundles from six different transgenic lines of OX5067 were completed, with no DsRed2 fluorescence detected. Each line represented a putative unique insertion event and was considered a replicate. Pseudoreplicates of males from the same line were also dissected to ensure against intra-line phenotypic variation.

Table 17. Summary of lines examined for DsRed2 positive sperm bundles and testes. No fluorescence was detected relative to a wild-type control.

| Putative line (OX5067) | Males examined | DsRed2-positive testes | DsRed2-positive sperm bundles |
|-----------------------------------|-----------------------|-----------------------------------|--|
| A1 | 3 | 0 | 0 |
| A2 | 1 | 0 | 0 |
| C3 | 2 | 0 | 0 |
| D1 | 1 | 0 | 0 |
| E1 | 1 | 0 | 0 |
| F1 | 1 | 0 | 0 |
| G2 | 1 | 0 | 0 |

Photographs were not taken for each replicate as Brownian motion and sperm movement combined with the long exposure time created blurring issues.

4.3.5.5 *B2-TUBULIN*-REGULATED DSRED2 AND PROTAMINE TRANSCRIPTION IN OX5067

Molecular detection of mRNA transcripts is likely to be more sensitive than visual detection of fluorescent proteins, if transcription and expression are low. Accordingly RNA was extracted from two random lines of OX5067. For each line, pupae were used, two males and two females from OX5067C1 and OX5067G2. A wild-type male and female were also included in the analysis. RNA was extracted (Methods 8.2.3).

This analysis may also have been useful in determining whether successful splicing of the *Dm* protamine is occurring, as if it was not, this would cause the DsRed2 transcript to be out of frame and therefore not translated correctly. The *Dm* protamine sequence includes an intron of 49 bp between protamine exons, facilitating differentiation between gDNA and ‘correctly’ spliced mRNA, where the protamine intron has been removed successfully (Figure 58).

| | |
|-----|---|
| 1 | GTACCGAAAA TTTGGGAAAA TTTGTTCTGT ATAAATAAG TGAGAGACTCT AGTGTATTA TAATTGAAAA ATGAGTCAA ATAATGAAA TGAGTCAG CATGGCTTT AAAGCCTTT AAACAAGACA TAATTTATC ACCTCTGAGA TCACAAATAT ATTAACTTT TACTCAAGTT TATTACATT ACTCACGTT |
| 101 | AGCCTGTGGA ATGGCATATAA TTCCATTCTC GCAAAGATG AAAGTCTAA AGGTCTACT GAAATGTGA ATCATCCAA GAGGAGAGCA CCTCAAAAT TCGGACACT TACCGTATTA AAGGTAAGA CGTTTCTAC TTCAGGATT TCCAGAGTA CTTAACACAT TAGTAGGTT CTCCCTCTGT GGAGTTTTA |
| 201 | GTAAGCCAAT AAAGTCTGT GCAAAGCCG GCGCAAAGGC AGCCTGTGCC AAGGCGACTC GGCCCAAGGT CAAGTGTGCA CCGAGTCAGA AGTGCAGCAA CATTCGGITA TTTCAGGACA CGTTTCGGCG CGGCTTCGCG TCGGACACGG TTCCGCTGAG CGGGTCTCCA GTTCACACGT GGCTCAGTCT CAACGTGTT |
| 301 | GCAGGGACCT GTCACTAACAA ACGCTTATTG GAATTTCGTG CGTTTCTCC GAAAGAAGCA CTGTGACTTG AAGCCGCAGG AATTAATTGC AGAGGCCGCT CGTCCCCTGGA CAGTGTGTT TGCGGATAAA CTTAAAGCAC GCAAAGAAGG CTTTCTCTGT GACACTGAAC TTGGGCTGCC ITAATTIAACG TCTCCGGCA |
| 401 | AAAGCGTGGG CCGAGCTCCC GGAGCATAGA AAGGATAGAT ACCGCGGAT GTACGTATT TTTTTTAT ATATAGAAGA CTATTAAATT GTCCCTTCAG TTTCGCAACC GGCCTCGAGGG CCTCGTATCTA TGGCGGCCCTA CCATGCTATAA AAAAATAAATGATACTCT GATAATTAAA CAGGAAGTC |
| 501 | GCATGCAAGG TCACCCCCAG TGAAAGCCAC AAGCGCCGC GGATTTGCAC CATGGCTICC TCCGAGAACG TCATCACCGA GTTCATGCGC TTCAAGGTGC CGTACGTICC AGTGGTGGTC ACTTGGCTG TTGCGGCTG CCTAAACGTG CTACCGGAGG AGGCTCTTG AGTAGTGGCT CAAGTACGCG AAGTTCCAG |
| 601 | GCATGGNAGGG CACCGTGAAC GGCCACGGAT TCGAGATCGA GGGCGAGGGC GAGGCCCGCC CCTACGAGGG CCACACACCC GTGAACCTGA AGGTGACCA CGTACCTICCC GTGGCAGCTG CGGGTGCTA AGCTCTAGCT CCCGCTCCCG CTCGGCGCGG GGATGCTCCC GTGTGTTGG CACTT 587 bp CACTGGT |
| 701 | GGGGGGGGGG CCTGGCTTCTG CCTGGGACAT CCTGTCCTCCC CAGTTCAGT ACGGCTCTAA GGTTGACGTG AAGCACCCCG CGACACAUCC CGACTACAG CCCCCCCCGGG GACGGGAAGC GGACCCCTGTA GGACAGGGGG GTCAAGGTCA TGCCGAGGGT CCACATGCACT TTCTGGGGC GCCTGTAGGG GCTGATGGT |
| 801 | AAGCTGTCTT TCCCCGAGGG CCTCAAGTGG GAGCGCGTGA TGAACCTCGA GGACGGCGGC GTGGCGACCC TGACCCAGGA CTCCCTCTG CAGGACGGCT TTCGACAGGA AGGGGCTCCC GAAGTTCACC CTCGCGCACT ACTTGAAGT CCTGCGCCCG CACCGCTGGC ACTGGGTCTT GAGGAGGGAC GTCCCTGCCA |
| 901 | GCTTCATCTA CAAGGTGAAG TTCACTGGCG TGAACCTCCC CTCCGACCGC CCCGTGATGC AGAAGAAGAC CATGGGTCTGG GAGGCTCCA CCGAGCCGCT CGAAGTAGAT GTTCACCTC AAGTAGCCG ACTTGAAGGG GAGGCTGCCG GGGCACTAGC TTCTCTCTG GTACCCGACCC CTCCGGAGGT GGCTCCGGGA |

Figure 58. The OX5067 Dm-protamine-DsRed2 transgene. The start codon is highlighted in green, the protamine exon sequences are highlighted yellow, the protamine intron sequence is grey and the DsRed2 coding sequence is red. The intron is 49 bp in length. Using primers specific for the first protamine exon and DsRed2 amplify the intron containing region in a PCR. Transcription and correct splicing of the intron would result in a sequence 557 bp in length.

RT-PCR targeting the transgene was conducted on RNA extracted from the male and female pupae. The initial attempt using two-step RT-PCR produced faint bands (not shown). A one-step no-RT control was conducted to mitigate RNA degradation between steps (Figure 59).

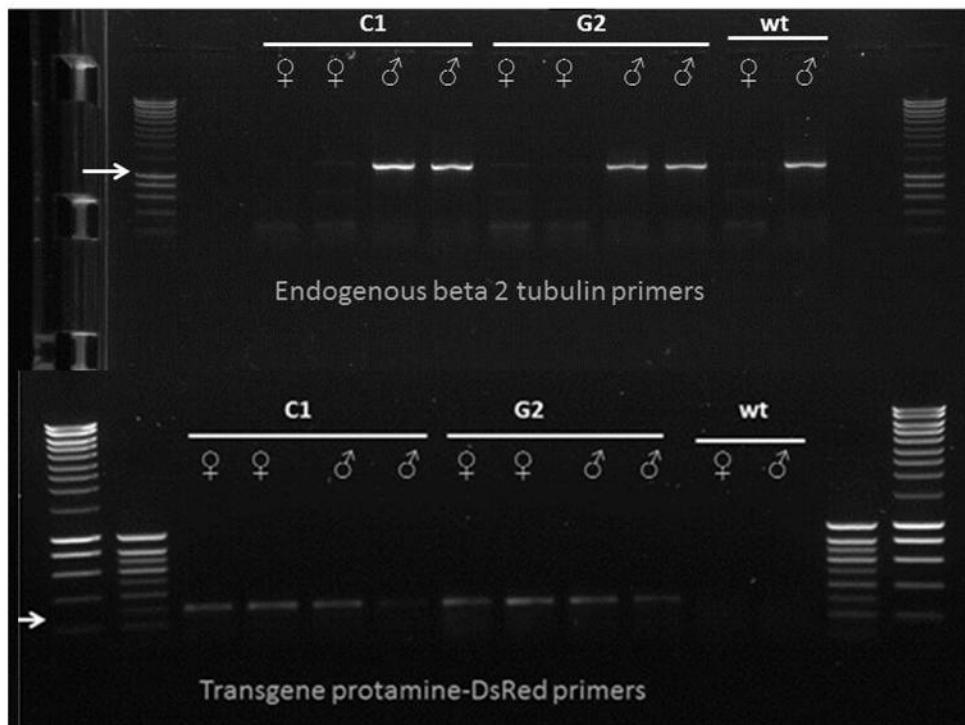


Figure 59. Molecular assessment of the presence of the transgene transcript in sex-specific fashion. A one-step RT-PCR produced from the RNA samples of OX5067 lines C1 and G2 and wild-type (wt) controls. Two male and female replicates from each line were used and one of each sex from the wild-type. The endogenous $\beta 2$ -tubulin primers GF2753 and GF2841 spanned the diamondback moth endogenous $\beta 2$ -tubulin gene. The primers in the second reaction type included TD825 and TD3275 amplifying from the first protamine exon to the start of the DsRed2 across the protamine intron and second exon. The endogenous $\beta 2$ -tubulin gene shows sex-specific transcription. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1 kbp, then by 500 bp to 3 kbp (Eurogentec, SF Smartladder). (Eurogentec, SF Smartladder). Inside ladder is 100 bp increments to 1 kbp.

The amplicons suggested that the transcript was present in the transgenic insects albeit at very low concentrations considering forty cycles of amplification. There is no evidence of differential transcription between the sexes, suggesting that the $\beta 2$ -tubulin promoter used with the transgene is not working in a similar fashion to the endogenous gene.

If splicing were correct we would expect a band at 342 bp. Sequencing the band of one male and one female will assist in determining if splicing is occurring in either sex, but this was deferred until sex-specific transcription was achieved.

4.3.5.6 A MORE CONSERVATIVE CONSTRUCT

It was noted that the first exon of $\beta 2\text{-tubulin}$ is small with the first intron close to the promoter.

It was deemed plausible that the first intron may have some enhancer activity in the transcription of the endogenous gene, as this characteristic can be seen in other species. This observation redirected the investigation towards the redesign of our construct, going back to basics, and removing the *Dm* protamine altogether. The first $\beta 2\text{-tubulin}$ intron was to be included in the construct upstream of the reporter DsRed2. This was supported by a re-reading of the insect $\beta 2\text{-tubulin}$ literature. Note 18 in Raff *et al.*, 1997 describes how:

"We obtained wild-type beta 2 tubulin-like levels of expression only with an intron-containing insert, suggesting that splicing may be important in normal beta 2 tubulin expression."

The intron-less insert of OX4703 and OX4804 did not result in testes-specific expression thus the new design included the first exon, first intron and truncated second exon (enough to ensure appropriate splicing). This went against historical success of $\beta 2\text{-tubulin}$ requisite sequence for the paternal effect system in the Diptera; but I was working with Lepidoptera.

4.3.6 OX5116 – IS INTRONIC DNA NECESSARY FOR EFFICIENT TRANSCRIPTION?

In building construct OX5116, only one copy of the SV40 sequence was successfully included on one side of the transgenesis marker module. It is regarded as bi-directional but embryos were inspected for transient expression following injection to confirm this. This was observed at a typical rate (Chapter 3), suggesting that the transformation marker module would be functional (see Appendix 9.1.7 for design and synthesis details).

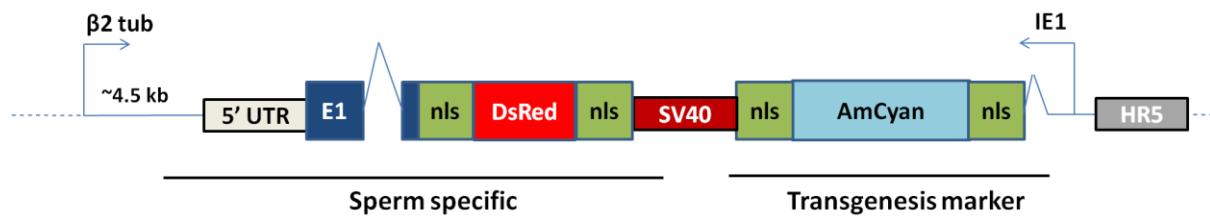


Figure 60. A simplified representation of the genetic construct OX5116, a more detailed description combined with component summary can be found in the Appendix 9.1. The main transgenesis marker is provided by the immediate early promoter (*HR5-*ie1**) with AmCyan-SV40. The sperm specific components were regulated by about 4.5 kb upstream of diamondback moth β 2-tubulin extended into the 5' UTR, first exon, intron and second exon. The DsRed2 was flanked by nuclear localisation signal (nls) to focus the expression in the head of the sperm. The dashed line represents other components of the circular construct omitted for clarity. The SV40 component is simplified, for this construct it was only on the 3' end of the DsRed2 not the AmCyan.

4.3.6.1 TRANSGENESIS

Embryos were injected and injection survivors reared and crossed (Methods 8.3.6.1).

Table 18. The injection data summary for the construct OX5116 in diamondback moth. Survival was 39%, above the median suggesting an improvement relative to OX4703 and OX5067 survival. Brackets provide survival and transformation efficiency.

| Species | Construct | Embryo injections (n1) | Injection survivors (G0) | Independent transgenic lines (X) |
|----------------------------|-----------|------------------------|--------------------------|----------------------------------|
| <i>Plutella xylostella</i> | 5116 | 1747 | 681 (0.389) | 2 (0.003) |

DNA was extracted from 5116A1 and 5116B2 adults and a PCR run targeting the sperm-specific transgene. The OX5116 gDNA was compared to OX5116 cDNA, OX4703A gDNA and wild-type gDNA (Figure 61).

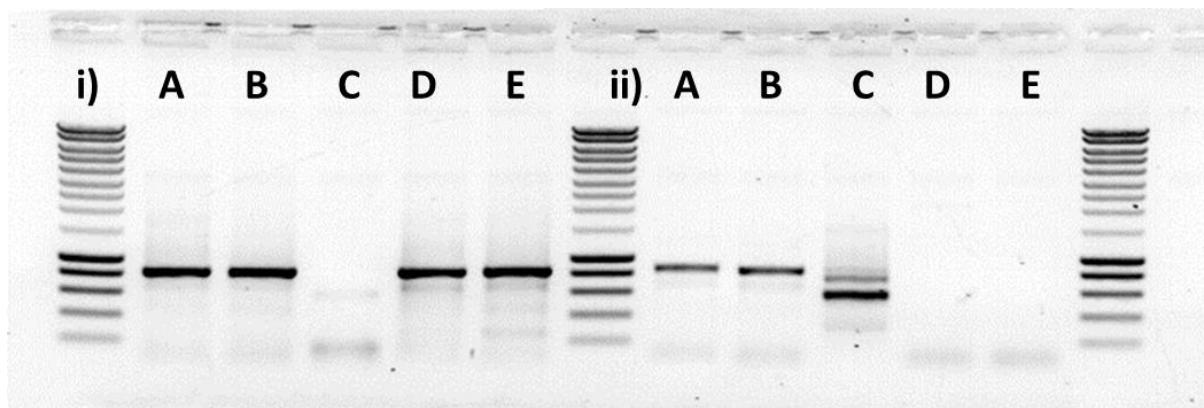


Figure 61. OX5116 DNA contains OX5116 specific sequence confirmed by PCR. Two PCRs were run on the putative transgenic lines to assess quality of the gDNA following extraction using primer pairs. The PCR product from reactions i) primers TD3637 and TD3638 (endogenous $\beta 2$ -tubulin specific, expected 771 bp) and ii) TD3658 and TD1406 (OX5116 transgene specific). 1 μ l of product was run with 2 μ l of 6x loading dye on a 1.2% agarose gel at 120 V for 30 minutes. Template nucleic acid included: A) OX5116A1 adult; B) OX5116B2 adult; C) OX5116 cDNA of the transgene; D) OX4703A gDNA; and E) wild-type gDNA. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1 kbp, then by 500 bp to 3 kbp (Eurogentec, SF Smartladder). (Eurogentec, SF Smartladder).

PCRs showed strong amplification with good specificity and expected amplicon sizes. A product was produced for each of the gDNA sampled as it amplified across the endogenous $\beta 2$ -tubulin. The transgene was detected only in OX5116 gDNA (and in the cDNA at the lower expected size, as introns were removed).

4.3.6.2 BACKCROSSES OF PUTATIVE LINES

Males were set up in crosses with wild-type females to perpetuate the lines. Transformants derived from at least two crosses, conservatively labelled as OX5116A1, OX5116A2, OX5116B1 and OX5116B2 (Table 19).

Table 19. Mendelian ratios of OX5116 and wild-type crosses. Putative transgenic G1 males derived from unique G₀ crosses were separated and labelled as different putative lines. One transgenic male was crossed with five wild-type females. Eggs were collected and reared to pupae. The frequency of transgenic to non-transgenic pupae were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios (or a proportion significantly greater than 50% transgenic (indicated by AmCyan fluorescent protein marker) to non-transgenic phenotype). Sex ratios were also inspected to rule out sex chromosome insertion of the transgene (W in the Lepidoptera, females are heterogametic). Line A2 was assumed sterile or died before mating.

| Line | Transgenic | Non-transgenic | X ² (2.d.p) | P(2.d.p) |
|------|------------|----------------|------------------------|----------|
| A1 | 86 | 68 | 2.10 | 0.15 |
| A2 | 0 | 0 | - | - |
| B1 | 38 | 39 | 0.01 | 0.91 |
| B2 | 22 | 32 | 1.85 | 0.17 |

The evidence suggests that the lines consist of a single autosome transgene insertion as transformants of both sexes were found.

4.3.6.3 TESTES DISSECTION AND FLUORESCENCE MICROSCOPY

4.3.6.3.1 PUPAE

Three male pupae from each line were dissected and testes removed. Testes were screened for DsRed2. All the testes screened were negative for DsRed2.

4.3.6.3.2 ADULTS

Adult testes were removed and screened under fluorescent microscope (x50 magnification). Six 5116A1 males were screened with none positive for DsRed2.

4.3.6.3 B2-TUBULIN-REGULATED DSRED2 TRANSCRIPTION IN OX5116

There is the possibility that the *β2-tubulin* fusion may have affected the DsRed2 from functioning as desired. Accordingly the transcription was assessed using one-step RT-PCR following RNA extraction (Figure 62).

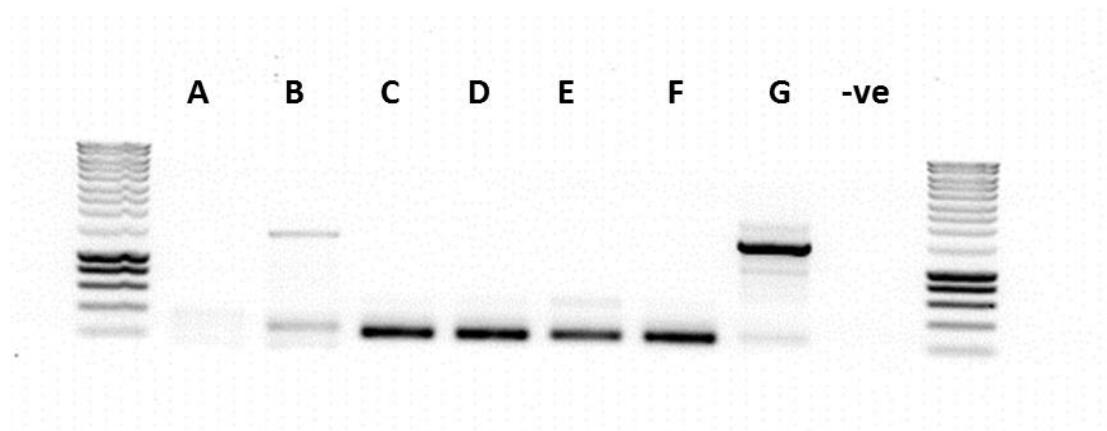


Figure 62. The $\beta 2$ -tubulin-regulated-transgene transcript was OX5116 testes-specific. Visualisation of the one-step RT-PCR product of extracted RNA from A) OX5116A1 male organs non-testes, B) OX5116A1 larvae third instar testes, C) two pairs of testes pupae OX5116A1, D) one pair of testes pupae OX5516A1, E) OX5116A1 adult testes, F) OX5116B2 adult testes, G) 4x wild-type pupae, -ve) water instead of template cDNA. The primer pair TD3658 and TD825 were used. A band of successfully spliced mRNA at 240 bp was expected in life stages C-F (unspliced at 477 bp). Sample B was too young for spermatogenesis. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

Interestingly the control of wild-type gDNA has shown some amplification with a 1500-2000-bp fragment, possibly the result of non-specific amplification as the forward primer was predicted to bind to the endogenous $\beta 2$ -tubulin gene as well as the version on the transgene. Each of the life stages of OX5116, from larvae to pupae show some product with negligible gDNA contamination. All OX5116 samples showed strong amplicons except for the RT-PCR on RNA extracted from a single pair of third instar testes. This is probably due to transcription of $\beta 2$ -tubulin starting after this life-stage.

Samples C and F were relatively clean so chosen for sequencing and comparison by alignment following purification. Both mRNA reads were compared to the OX5116 sequence to assess whether splicing was correct. The correctly spliced transcript was produced with DsRed2 in frame. Despite apparently abundant transcript, no DsRed2 fluorescence was detected, possibly because the first and second exon (24 aa) were interfering with tetramisation. A cysteine amino acid may have been responsible as the rest are mainly glycine.

A summary of the sequence and expected translation are given below (Figure 63).

3658-B2TubKpnF 100.0%

11901 AACACACAAT ATTTGAATGA GGAATGTTTT TCAAGTCAC TCATTTTTT TCCCTGTAG TAGGTGAG TGACCTATC AGGGAGGAA CGAATATT
3658-B2TubKpnF 100.0%
+3 Met Arg Glu Ile Val His Ile Gin Ala Gly Gin Cys Gly Asn Gin

12001 GCGAAATTTG TTCTCTTAITA AATAAAGTGGA GACTCTAGTGT TTATTATATAT TGAAAATATGC CTCAAATTTG TCACATACAA GCCTGGCGGT CGCGGACCA
+3 Gln Ile Gly Ala Lys

12101 GATGGAGCT ARGGTAAATT TCATTTCTAC ATACATTTT ATTGAATCTC TGTCCTTTG CATAATTCA TGTTAACCTC ACGTAGGGTA AGCCCACGGA
12201 TCCATTAATCT TGCAGTCCTGT AGTGGCTAGA CAGGAACAA TAGGTATAGT GTTGCGCCGC TCCCTGATAG AAGCTCTGTA ACTGCAGGCA ATCTTCGTT
+3 Phe Trp Glu Val Ile Ser Ala Ser Glu Asn Val Ile Thr Glu Phe Met

12301 CGCTGCGCTA ATCACAAAGTT CTCTCTAATG ATTICCAATGT TACATTTCAAGT Phe Trp Glu Val Ile Ser Ala Ser Glu Asn Val Ile Thr Glu Phe Met
+3 Met Arg Phe Lys Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu Gly Glu Arg Pro Tyr Glu Gly His Asn Thr Val Lys

12401 GCGCTTCAAG GTGCCCATGG AGGGCACCGT GAACGGCCAC GAGTTCGAGA TCGAGGGCGA GGGCGAGGGC CGGCCCTACG AGGGCCACAA CACCGTGAAG
825) DsRed-DsRed 100.0%
+3 Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Glu Phe Glu Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile

12501 CTGAAGGTCA CCAAGGGCGG CCCCTGCCCC TTGCGCTGGG ACATCTGTC CCCCCCTTCAGT CAGTACGGCT CCAAGGGTCA CGTCAAGCAC CCCGCCGACA
+3 Ile Pro Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Ala Thr Val Thr Gin Asp Ser Ser

12601 TCCCCGACTA CAAGAAGCTG TCCCTCCCCG AGGGCTTCAA GTGGGAGCGC GTGATGAAC TCGAGGGACGG CGCGCTGGCG ACCGTGACCC AGGACTCCTC
+3 Ser Leu Gin Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn Phe Pro Ser Asp Gly Pro Val Met Glu Lys Lys Thr Met Glu Trp Glu Ala

12701 CCTGCAGGAC GCCTGCTTCA TCTACAGGT GNAGTTCATC GGCGTGAAC TCCCCCTCCGA CGGCCCCGTG ATGCAGAGA AGACCATGGG CTGGGAGGCC
+3 Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu Thr His Lys Ala Leu Lys Asp Gly Gly His Tyr Leu Val Glu Phe Lys Ser

12801 TCCACCGAGC GCCTGACCC CGCGCACCGGC GTGCTGAAGG CGGAGACCCA CAAGGCGCTG AAGCTGAAGG AGCGGCCCA CTACCTGGT GAGTCTAGT
+3 Ser Ile Tyr Met Ala Lys Lys Pro Val Glu Leu Pro Gly Tyr Tyr Val Asp Ala Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile Val Glu

12901 CCATCTACAT GCCAAGAAG CCCGTGCAGC TCCCCGGCTA CTACTACGTG GACGCCAAGC TGGACATCAC CTCCCCAACAC GAGGACTACA CCATCGTGG
+3 Glu Gin Tyr Glu Arg Thr Glu Glu Arg His His Leu Phe Leu Arg Ser Arg Pro Lys Lys Lys Arg Lys Val Glu Asp Pro ...

13001 GCAGTACGAG CGCACCGAGG GCCGCCACCA CCTGTCTCTG AGATCTGCAC CCAAGAAAAA CGCGGAAGGTG GAGGACCCGT AAGATCCACC GGATCTAGAT

Figure 63. Part of the construct OX5116 intended to cause sperm-specific expression of DsRed2. Sequencing revealed that the intron was successfully removed (grey). The nucleotide sequence of DsRed2 (red) appears to be in frame relative to the start codon (purple) without any stop codons in frame. The exons of β 2-tubulin (blue) may interfere with correct folding and tetramisation of DsRed2. The primers TD3658 and 825 are shown for reference.

4.3.6.5 SEMI-QUANTITATIVE TRANSGENE TRANSCRIPTION COMPARED TO THE ENDOGENOUS GENE

The RT-PCR method can be used not only to detect specific mRNAs but also to quantify their relative levels. This was achieved by comparison of the transgene with the endogenous β -tubulin gene (Figure 64).

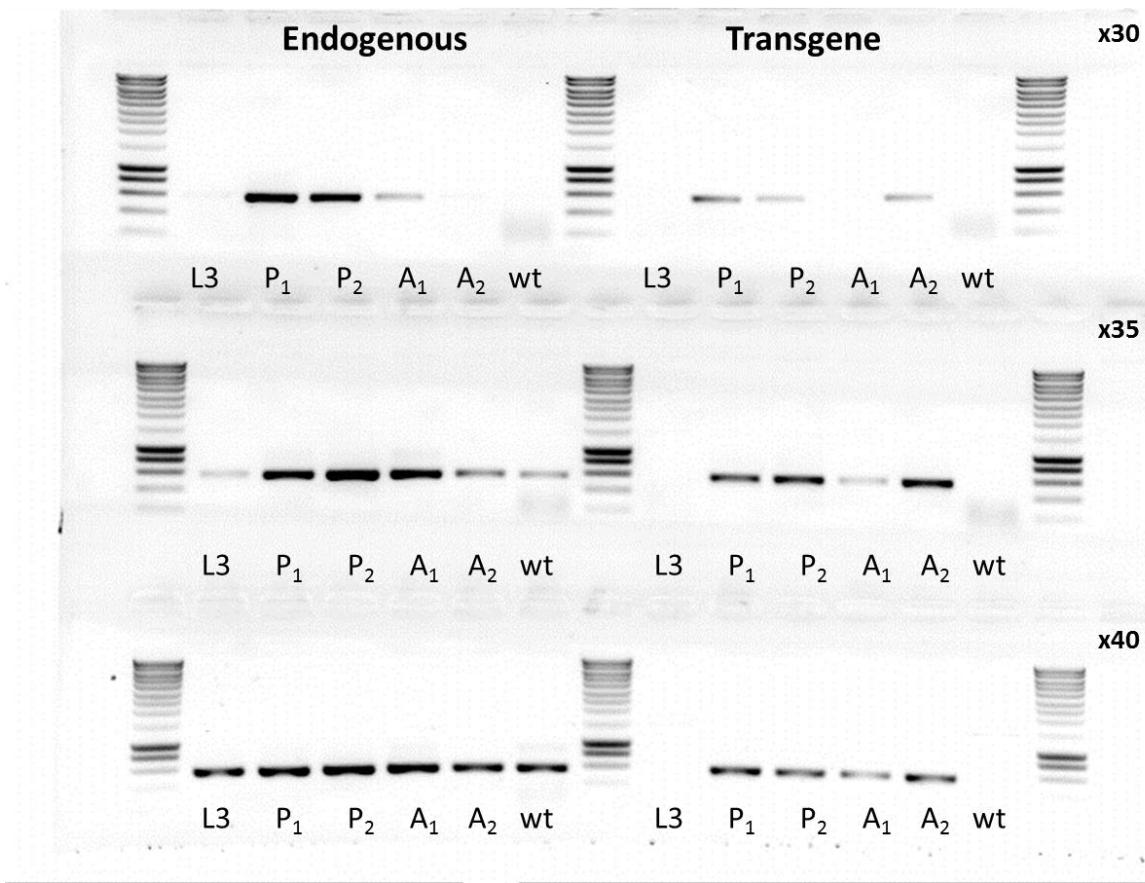


Figure 64. Transcription of the transgene closely mirrors that of the endogenous $\beta 2$ -tubulin, albeit showing slightly weaker transcription levels. A semi-quantitative one-step RT-PCR of RNA extracted from testes of: L3) OX5116A1 larvae, P₁) pupae OX5116A1, P₂) pupae OX5116B2, A₁) adult testes OX5116A1, A₂) adult testes OX5116B2, wt) whole male pupa. The reactions in the left column marked endogenous primers specific to the endogenous $\beta 2$ -tubulin gene (TD3637 & 3638). The right column is reactions specific to DNA with the transgene sequence of OX5116 across an intron to aid discrimination between cDNA and gDNA amplification (TD3658 & TD1406). Both sequences are expected to have a band size of 574 bp. The RNA was added to the reaction at 0.0595 μ g in 20 μ l with standard methodology except for pausing of thermocycler at desired cycle number for extraction of 4 μ l of PCR product to be run on a gel. The null hypothesis of no difference was tested using a one-step RT-PCR with removal of product during the exponential cycles (PCR product removed at cycles 30 and 35). The wild-type sample was negative for the transgene PCR. Smartladder is included in the middle and side wells, in the bottom row (x40) the 200 bp rung of the ladder is not visible. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

The endogenous gene was amplified to saturation for all samples by the 40th cycle. The larvae had the lowest amount of endogenous transcript. This may be due to gene regulation; with activation of $\beta 2$ -tubulin transcription coinciding with spermatogenesis. This corroborates the transcriptome data from Liverpool, which suggests that the gene is not transcribed until late larvae. Transcript levels appear highest in the pupae with saturation of the product at 30 cycles compared to the adults which required 40 cycles for saturation.

This relative trend by life stage was seen in the transgene also, with relative abundance ordered pupa, adult then larva (high to low). However, comparison between genes suggests the transgene transcript was not at as high levels as the endogenous gene transcript. With both replicates forming bands at 35 cycles.

Although reduced the transgene transcription is assumed to be at acceptable levels for use of the same promoter, 5' UTR, exon 1, intron 1 and exon 2 sequences, for use in a future construct. The lower transcription rate may be due to chromatin effects.

4.3.6.6 SPECIFICITY OF THE TRANSGENE

The levels of the transcript have been assessed in the testes of various life stages which conform to the expectation based on the endogenous gene regulation. However, it has not been confirmed that transcription does not occur in the females or male non-testes tissues at detectable levels. Accordingly RNA was extracted from appropriate samples and a one-step RT-PCR conducted. An endogenous control of *doublesex* primers was used to assess the quality of the RNA.

A female whole carcass, male testes and the upper or anterior half of the male carcass were dissected. The RNA was extracted from OX5116A1, OX5516B1 and wild-type insects. The RNA was acceptable with no extreme degradation. A one-step RT-PCR for 40 cycles was conducted on the RNA with two distinct primer pairs. Given the design it was expected that only the OX5116 testes samples would produce the 225 bp amplicon (Figure 65).

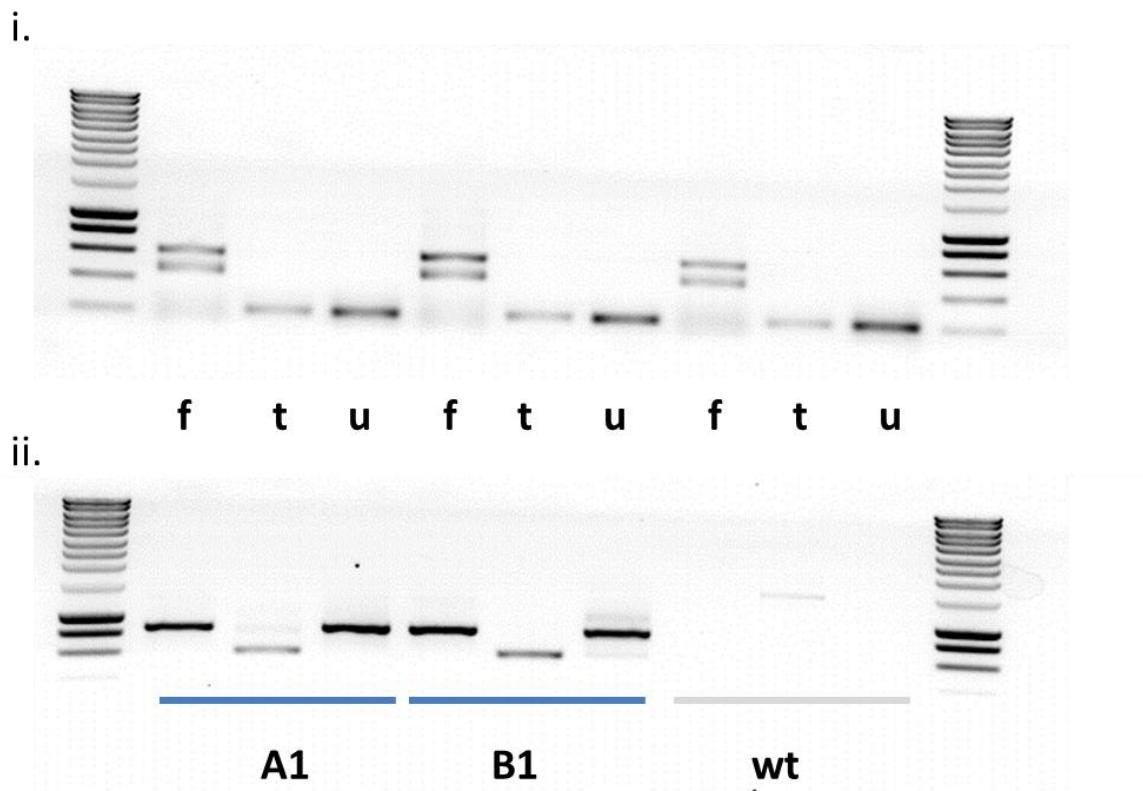


Figure 65. OX5116 demonstrates testes-specific transcription of a transgene requires the $\beta 2$ -tubulin promoter, 5' UTR and some additional sequence beyond the start codon, possibly the first intron. One-step RT-PCR products were run in an agarose gel using default methods. Two reactions were conducted on each sample; i) diamondback moth *dsx* primers TD3482 (F primer in exon 2) and TD3544 with expected band size in the female 468 bp and male 225 bp. The endogenous control simultaneously assessed the quality of the RNA used in the process and the sex of the insect used. ii) The OX5116 transgene-specific primers used in reaction two were TD3658 and TD1406 with expected correctly spliced mRNA band size of 574 bp. The RNA was extracted from OX5116A1, OX5516B1 and wild-type insects. A female whole carcass (f), male testes (t) and the upper half of the insect left over after testes were dissected (u) from each group that was used. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

The banding was as expected for the endogenous control with a female band at 468 bp and male at 225 bp. However, an additional band near to the 600 bp ladder band was unexpected and hitherto unseen in previous reactions carried out with this primer set, it may be a different female isoform. The testes amplicon is fainter than that from the upper half of the insect. For the transgene reaction only the OX5116 testes samples were positive for the transgene transcript. Combined with the previous semi-quantitative RT-PCR this data suggests the revised $\beta 2$ -tubulin promoter and non-coding DNA provide appropriate specificity and transcription.

4.3.6.7 OX5116 TRANSGENE PROMOTER PROVIDED TRANSCRIPTION BUT NOT TRANSLATION

Transcription appears to be functioning as designed under control of the *β2-tubulin* sequence used in OX5116, yet no DsRed2 fluorescence is visible in the testes. This is unusual in some eukaryotes as typically transcription — not translation — determines protein abundance (Li & Biggins, 2015).

DsRed2 fluoresces green as a monomer and then red after tetramisation. The first exon and part of the second exon of *β2-tubulin* lay upstream of DsRed2 in our transgene mRNA. At 24 aa in length, this may interfere with protein-folding or tetramisation. Accordingly when screening we checked for DsRed2 using both red and green filter conditions. No fluorescent sperm were detected.

4.3.7 CREATING A CONSTRUCT WITH THE *B2-TUBULIN* AND DSRED2 TRANSGENE SEPARATED BY UBIQUITIN

Ubiquitin (the acronym Ub refers to both free ubiquitin and the ubiquityl moiety) is highly conserved in eukaryotes and has a variety of roles. In designing a new construct, Ub coding sequence was included to act as a spacer with the eukaryotic cellular machinery recognising it and cleaving the *β2-tubulin* exons in the transgene from the DsRed2. This artificial principle role of Ub is reviewed in Schnell and Hicke (2003).

Varchesky (2005) describes the ubiquitin fusion technique, whereby linear fusions of Ub to other proteins are rapidly cleaved by endogenous proteases after the last residue of Ub following translation. The proteases involved are called deubiquitylating enzymes (DUBs). Cleavage takes place regardless of the identity of the junctional amino acid residue. The junctional amino acid can affect the half-life of the co-translated protein.

The new construct was modified from OX5116 by the addition of the ubiquitin in between the transgene $\beta 2$ -tubulin and DsRed2, designed to result in co-translation followed by cleavage of the two proteins. The first amino acid in DsRed2 was compatible with the technique, predicted to create a protein with a long half-life (Varshavky, 2005).

It was desirable to have a protamine-DsRed2 fusion to ascertain whether *Dm* protamine would serve as a means to deliver a protein to the DNA in the nucleus of a diamondback moth sperm cell (Appendix 9.1.9).

4.3.8 OX5133 AND OX5135

Protamine coding sequence was added upstream of the DsRed2 coding sequence of OX5133 to assist in chaperoning the DsRed2 to the nucleus of the sperm, in anticipation of the paternal effect system requiring a mechanism to deploy a nuclease to the sperm DNA and create double-stranded breaks (Figure 66).



Figure 66. The OX5135 simplified schematic is similar to OX5116 with the addition of ubiquitin and *Dm* protamine between DsRed2 and the $\beta 2$ -tubulin region. The $\beta 2$ -tubulin promoter (arrow) and 5' UTR lay upstream of the first exon (E1), intron (bent line), second exon (E2), ubiquitin (Ubi), *Drosophila melanogaster* protamine (Prot) and DsRed2.

 4.3.8.1 TRANSGENESIS

 4.3.8.1.1 OX5133- 'IN UBIQUITIN WE TRUST'

Table 20. Injection data summary for the construct OX5133 in diamondback moth. Brackets provide proportions for survival and transformation efficiency.

| Species | Construct | Embryo injections (n_1) | Injection survivors (G_0) | Independent transgenic lines (X) |
|----------------------------|------------------|---|---|---|
| <i>Plutella xylostella</i> | OX5133 | 1934 | 466 (0.24) | 2 (0.004) |

DNA was purified from 5133A1 and 5133A2 adults and genotyped by PCR to assess the presence of the transgene. The PCRs failed, possibly due to the high GC content of the newly added ubiquitin region. Appropriate primers and *Taq* polymerases were explored further (4.3.8.4 Sequence check of transgene region).

4.3.8.1.2 OX5135

Table 21. Injection data summary for the construct OX5135 in diamondback moth. Only one egg collection was completed before the crosses were disposed of accidentally. Brackets provide proportions for survival and transformation efficiency.

| Species | Construct | Embryo injections (n_1) | Injection survivors (G_0) | Independent transgenic lines (X) |
|----------------------------|-----------|-----------------------------|-------------------------------|----------------------------------|
| <i>Plutella xylostella</i> | 5135 | 1351 | 354 (0.26) | 0* (NA) |

OX5133 crosses were prematurely discarded by accident.

4.3.8.2 BACK-CROSSES OF OX5133 PUTATIVE LINES

Crosses with wild-type were set up with one transgenic female to two males or one transgenic male to five wild-type females. The second set of lines — designated 'B' — were lost before they could be assessed.

Table 22. Putative transgenic G_1 insects derived from unique G_0 crosses were separated and labelled as different putative lines. If male it was crossed with five wild-type females, if female two males. Eggs were collected and reared to pupae. The frequency of transgenic to non-transgenic pupae were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios (or a proportion significantly greater than 50% transgenic (indicated by AmCyan fluorescent protein marker) to non-transgenic phenotype). Sex ratios were also inspected to rule out sex chromosome insertion of the transgene (W in the Lepidoptera, females are heterogametic). Line A3 had a non-Mendelian distribution possibly due to negative effects of the transgene on the insect or poor penetrance and or marking. Proportions of the total are given in brackets.

| Line | Transgenic | Non-transgenic | χ^2 (2.d.p) | P (2.d.p) |
|------|------------|----------------|------------------|-----------|
| A1 | 49 (0.45) | 59 (0.55) | 0.93 | 0.34 |
| A2 | 37 (0.49) | 39 (0.51) | 0.05 | 0.82 |
| A3 | 49 (0.39) | 76 (0.61) | 5.83 | 0.02 |

4.3.8.3 TESTES DISSECTION AND FLUORESCENCE

4.3.8.3.1 PUPAE

From each 'A' line, four male testes were dissected and screened for DsRed2 using fluorescence microscopy, none was detected. Microphotographs of non-fluorescent testes or sperm are not shown.

4.3.8.3.2 ADULTS

From each 'A' line, three male testes were dissected and screened for DsRed2 fluorescence microscopy, none was detected.

4.3.8.4 SEQUENCE CHECK OF TRANSGENE REGION

DNA was extracted from pupa from two of the putative lines, OX5133A1 and OX5133A2. A PCR was run with different primers to previous constructs due to the putative problems caused by the high GC content of the ubiquitin region newly added to this construct.

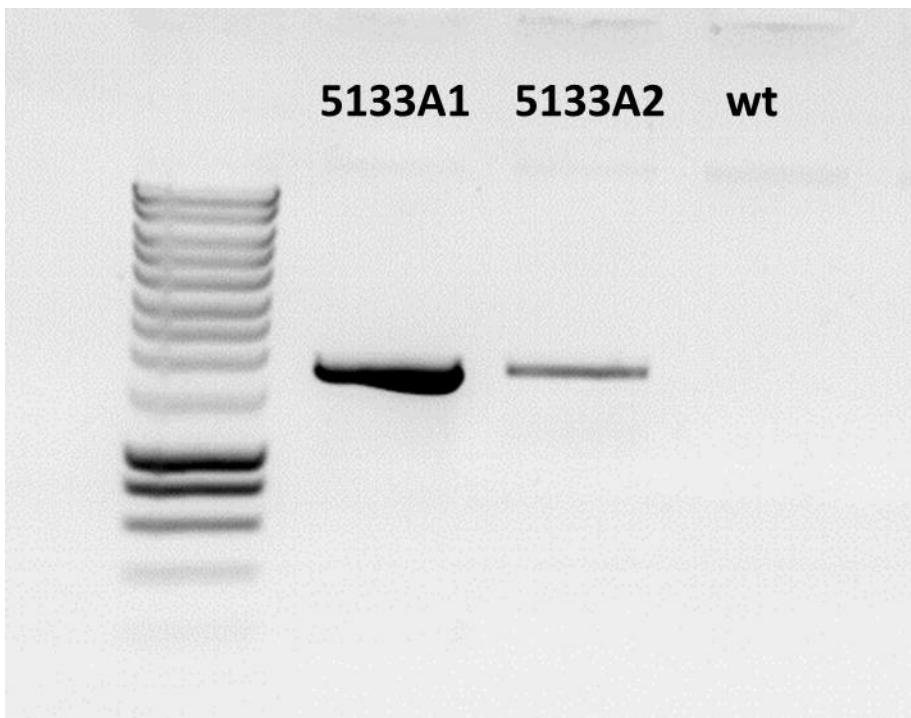


Figure 67. OX5133 construct-specific DNA detected in putative lines. The PCR product from reactions with primers TD3637 and TD1128 (transgene specific) using Herculase at 30 cycles and using recommended reagent concentrations and volumes. 1 μ l of product was run with 2 μ l of 6x loading dye on a 1.2% agarose gel at 120 V for 30 min. Template nucleic acid included: OX5133A1, A2 and wild-type (wt) gDNA. A band size of 1697 bp was expected. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The amplicons were checked on a gel (Figure 67), column-purified and then sent for sequencing with primers targeting the DsRed2 coding region. The primers were designed such that transgene of interest was sequenced in both directions (TD3057, 932 bp; TD38, 863 bp).

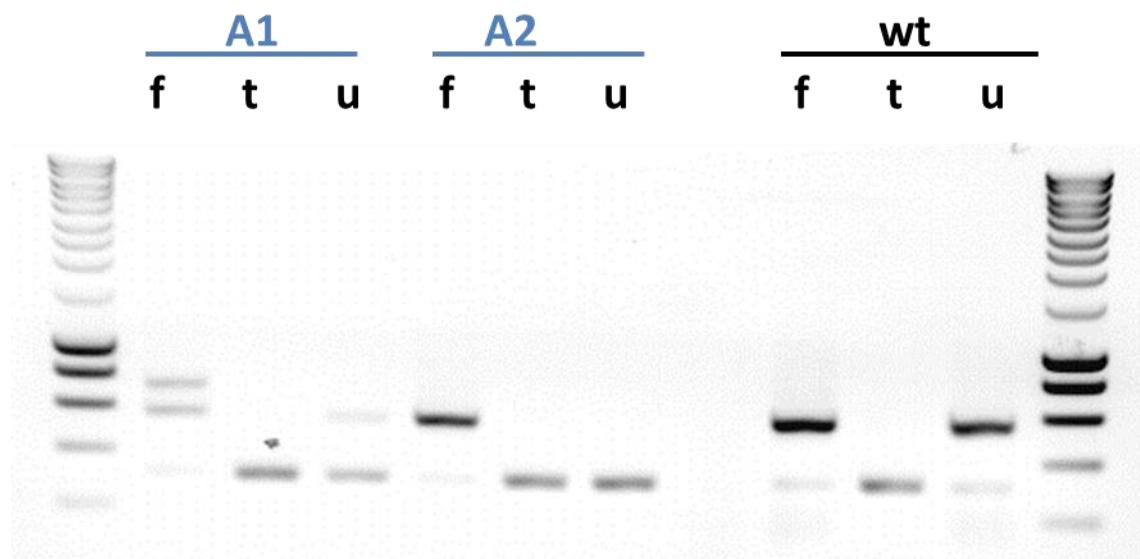
In total there were four reactions, two for each putative line. Sequences were aligned and compared using EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html). Both reads were identical. Some nucleotides (10-20 bp) were missing at both ends due to non-identical coverage (an 8-nucleotide overlap was in the middle of the pair of sequences). The sequence data was identical to our design of OX5133.

Comparison of the exon 2 in construct OX5133 with OX5116 reveals the same sequence adjacent to the intron (TTCTGGAAAGTGATATCG) ruling out any sequence differences adjacent to the splice site.

4.3.8.5 *B2-TUBULIN* PROMOTER-REGULATED DSRED2 TRANSCRIPTION IN OX5133

The specificity of the transcription was to be determined using RT-PCR (Figure 68) using alternative primers to those used when examining OX5116 [these primers failed (not shown) possibly due to the ubiquitin increasing the target size and making the target transcript more susceptible to degradation].

Endogenous



Transgene

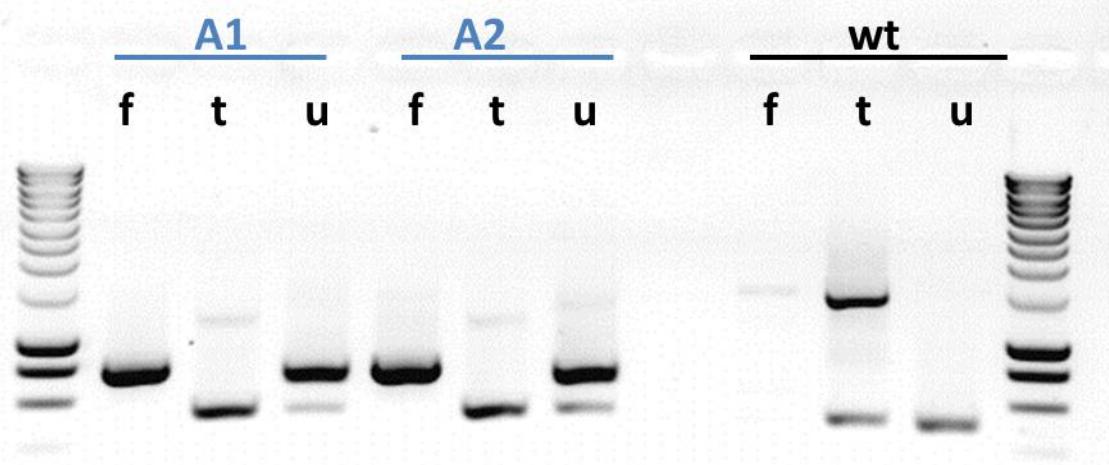


Figure 68. Testes-specific transcription of the transgene in OX5133 positive insects. One-step RT-PCR products were run in an agarose gel using default methods. Endogenous primers targeting the β -tubulin gene were expected to give a band size of 313 bp for mRNA (gDNA 550 bp) and OX5133 transgene-specific primers TD3637 and TD825 with expected correctly spliced mRNA were expected to give a band size of 384 bp (gDNA 546 bp). The RNA was extracted from OX5116A1, OX5516B1 and wild-type insects. A female whole carcass (f), male testes (t) and the upper half of the insect left over after testes were dissected (u) from each group that was used. The transgene performed well giving male-specific transcription. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

The wild-type control showed no transgene transcript of the expected size with a band slightly smaller than the expected 384 bp. The wild-type showed the strongest banding for the endogenous transcript. Targeting the tubulin may be a poor endogenous control as due to the similarity amongst the tubulin genes non-specific banding is seen.

The transgene transcript was not present in females but in males, and more strongly expressed in the testes than in the male carcass upper half. RT-PCR will not necessarily indicate whether the ubiquitin is functioning as desired, as it causes amino acid chain separation at the translation step.

The amplicon bands of OX5133A1t and OX5133A1u at 200-400 bp (Figure 68) were gel-purified, cloned and sequenced (not shown). A representative transcript sequence from the testes was compared to the expected correctly spliced mRNA derived from the electronic version in Vector NTI. The sequences were compared with most of the transgene covered by the sequencing except for regions of the *β2-tubulin* exon one and the 5' UTR (Figure 69). The ATG start codon is found 121 bp into the first exon (the *β2-tubulin* exons are merely included in the design facilitating transcription). The nucleotide sequences of the features were compared systematically with up to one mismatch allowed per 20 bp. If translation were to begin at the intended methionine (ATG) the DsRed2 should have been in frame.

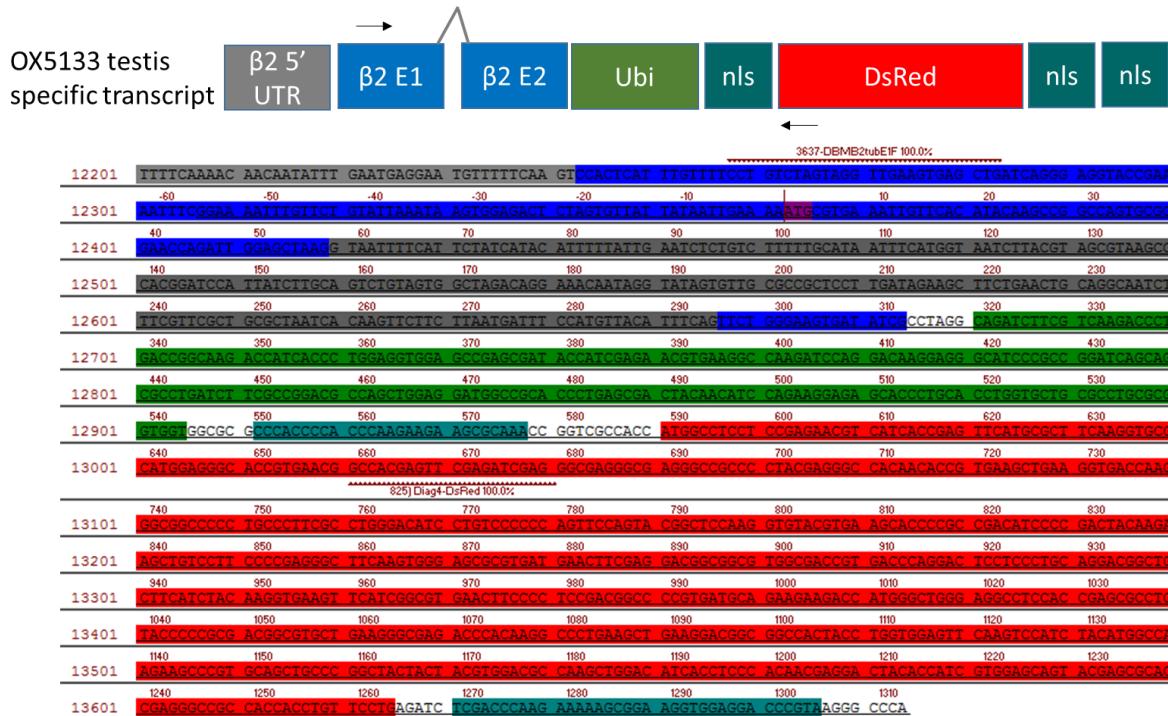


Figure 69. Schematic of the OX5133 compared to actual sequence in OX5133 positive insects. The transcript sequence derived from mRNA extracted from OX5133A1 testes is shown as the underlined nucleotides compared to the designed sequence of OX5133. The coding sequence of the transgene was as expected. The sequencing coverage was from the end of the $\beta 2$ -tubulin exon one ($\beta 2$ E1 - 177 bp) to the SV40 region at the 3' end of the transgene, including complete sequences for the $\beta 2$ -tubulin exon two ($\beta 2$ E2 - 18 bp), ubiquitin (Ubi - 225 bp), nuclear localisation signal (nls - 27, 36 & 37 bp from left to right). A coloured key is given above the sequence. The white gaps in between the coding sequences were intentional and are made up of neutral and or small amino acids. The nucleotide number relative to the position in the construct is shown on the left. The red numbers above the direct strand of DNA shown represent relative position to the start site. The DsRed2 is in-frame. The band sequence from the carcass is not shown as it was found to be non-specific to the transgene. The intron between the exons is not spliced and may contribute to DsRed2 mis-folding. The SV40 at the 3' UTR is not shown.

OX5116 produced correctly spliced transgene, as did OX5133, yet we were not able to visualise DsRed2 fluorescence despite the addition of ubiquitin which if transcribed should provide cleavage and functional DsRed2. This suggests that the transcript may not be getting translated due to lack of some regulatory information. The endogenous $\beta 2$ -tubulin is correctly translated (assumed) typically concomitantly with transcription (at least in *Drosophila* this has been shown); our transgene contains the majority of the non-coding and some of the coding region except for most of exon 2 and the 3' UTR. The SV40 used in our constructs may lack the necessary “cross-talk” with the 5' UTR and other regulatory mechanisms in the germline which facilitate appropriate timing of translation of the genes transcribed prior to meiotic arrest. Following an inconclusive experimental assay of DsRed2 protein presence using molecular

methods (Methods 8.2.22), we proceeded with a new design incorporating the 3' UTR of the diamondback moth *β2-tubulin*.

4.3.8.6 HOW TO PROCEED GIVEN TRANSCRIPTION BUT NO TRANSLATION

Given the issue of no fluorescence due to assumed non-translation of the transcript, it was noted that with sperm-specific genes, sometimes the 3' UTR facilitates cross talk with other regulatory regions of the transcript and permits translation (although in *Drosophila*, SV40 3' UTR has been shown to consistently work with the *β2-tubulin* promoter driving transgene transcription and translation; Barckmann *et al.*, 2013). For example, in *D. melanogaster* transcripts for genes known to be required specifically during spermatogenesis, including for example the sperm tail protein Don Juan, are synthesised in the primary spermatocyte nuclei, and stored in a translationally repressed state for several days, until late in spermatogenesis (White-Cooper, 2010). Perhaps the 3' UTR is involved in timing the translation of these transcripts?

That we see endogenous gene translation but not transgene translation, despite transcription, may be due to the omission of the 3' UTR. A construct was designed to remedy this with successful translation to be tested by crossing with the reporter line OX4026 (Figure 70).

Proposed design

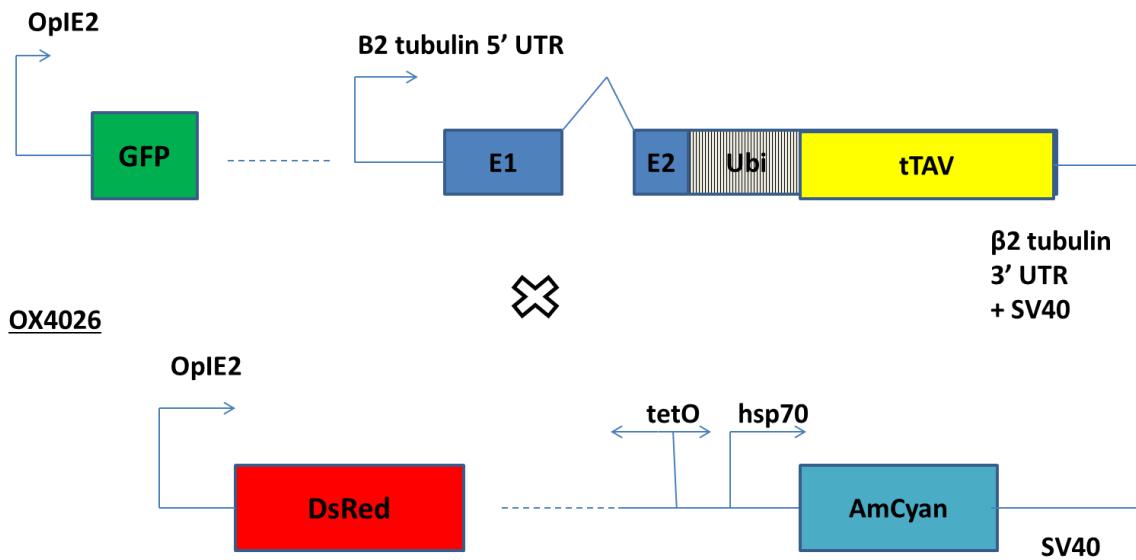


Figure 70. The experimental method to test whether the $\beta 2$ -tubulin 3' UTR is required for translation of the $\beta 2$ -tubulin-transgene. Crossing transformants of each line will result in cyan-fluorescent sperm in those males carrying the GFP (ZsGreen) and DsRed2 transgenes. This should be assessed by fluorescence microscopy with suitable controls for GFP and DsRed2 somatic bleed-through in the sperm.

4.3.9 DIAMONDBACK MOTH *B2-TUBULIN* AXENOME AND 3' UTR

Previously I had focused on the 5' region of the gene to capture the promoter and ensure transcription. Hitherto I had not conducted any RT-PCR on the 3' end of the transcript.

4.3.9.1 PUTATIVE 3' UTR UTILITY

The biological fate of each mRNA and consequently, the protein to be synthesised, is highly dependent on the nature of the 3' untranslated region. Despite its non-coding character, the 3' UTR may affect the final mRNA stability, the localisation, the export from the nucleus and the translation efficiency (Michalova *et al.*, 2013). As I experienced problems with translation of $\beta 2$ -tubulin-DsRed2 transcript I decided to investigate the role of the $\beta 2$ -tubulin 3' UTR to supplement the SV40 already employed.

With recent updates in the gDNA and transcriptome data for diamondback moth I proceeded to inspect the 3' UTR and design primers to amplify this region.

4.3.9.2 DIAMONDBACK MOTH *B2-TUBULIN* SEQUENCE IMPROVEMENT

Two diamondback moth sequences to consider:

- The original Liverpool sequence from which OX5133 was designed and derived from the RT-PCR corrected 7/1/2013 sequence. In addition to this 5' RACE was used to confirm the 5' UTR of the original sequence data and 4.5 kb of the upstream region was also cloned for inclusion in the construct. This was the more trusted data.
- Our new updated Liverpool scaffold data which was sampled from our insects, 1/2/2015 sequence data.

An alignment of the two sequences revealed great similarity with two noticeable exceptions, the newer sequence had a 30 bp deletion at nucleotide position 754-782. This deletion was located 378 bp upstream of the start codon in the promoter region of the gene. Another deletion of 40 bp at nucleotide position 1249-1290 was found in the intron between exons 1 and 2, near the 5' splice site (Figure 71).

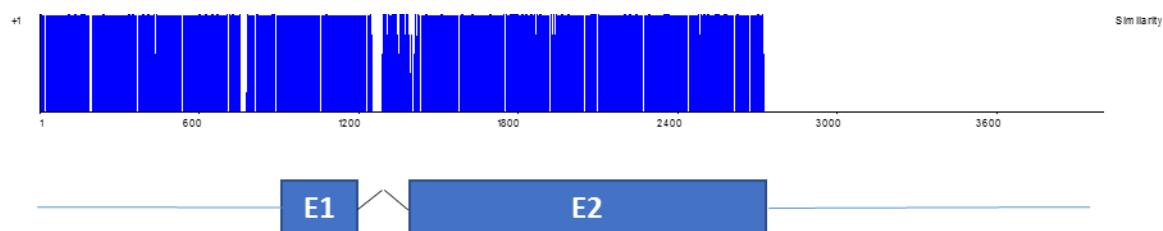


Figure 71. The similarity of the sequences is approximately aligned with a diagram of the diamondback moth *β2-tubulin* gene to show where the deletions occur in the newer sequence. The similarity is on the y-axis and uses default values. The x-axis shows the relative nucleotide position or number. The new Liverpool data provides additional sequence at the 3' end of the gene as shown by the zero similarity from 2721 onwards.

Accordingly one would expect no change in the coding sequence as only the untranslated regions experienced major changes. This was checked by translation and comparison of both sequences from the first ATG open reading frame in exon 1 (starting MREIV).



Figure 72. Amino acid sequence comparison revealed a frame shift caused by a deletion. The similarity of the amino acid sequences was compared using Vector NTI. The sequences were identical up to an amino acid deletion in the new Liverpool at position 99 shown by the sheer decline to zero similarity thereafter.

The first 98 amino acids were identical with a deviation occurring at amino acid 99. Inspection at the nucleotide level revealed a deletion of G at nucleotide 293 of the ORF as responsible for pulling the sequence out of frame and causing the discrepancy in amino acid sequence and the premature stop codon at amino acid 117 (Table 23). The RT-PCR data shows that this one nucleotide indel can be ignored and that the ORF is correct.

Table 23. A comparison of the newly acquired diamondback moth shotgun sequence data (December, 2014), provided by researchers at Liverpool, compared to our own sequence data. Our data was derived from RT-PCR and transcript sequencing of the diamondback moth $\beta 2$ -tubulin gene. The sequence shown is that from nucleotide position 410 (relative to exon 1 and exon 2 with the intron removed or the coding domain) to 434. The deletion is highlighted using an underscore in the new data to represent its absence and capital letter in the old sequence to represent its presence.

| Information source | DNA putative sequence |
|-------------------------------|--------------------------|
| Old sequence data from RT-PCR | ggGaataactgggctaaaggacac |
| New Liverpool sequence data | gg_aataactgggctaaaggacac |

The $\beta 2$ -tubulin 3' UTR was shown to extend approximately 77 bp beyond the stop codon (Figure 73), elucidated by web-based searches and tools (Methods 8.2.27). This may be a seemingly small sequence to be involved in translation regulation of the transcript.

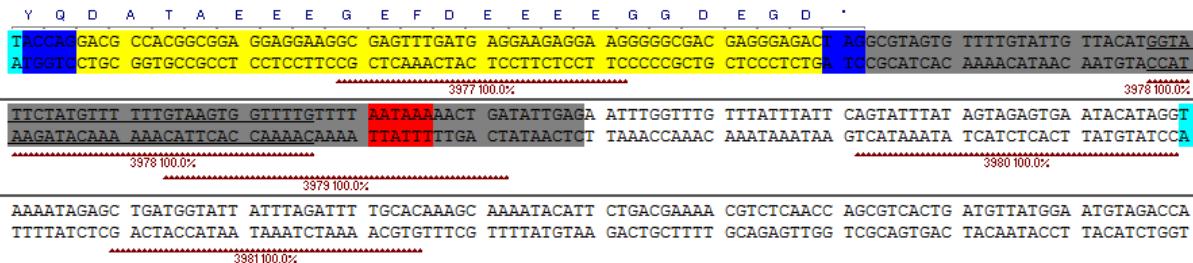


Figure 73. Updated diamondback moth *β2-tubulin* DNA sequence data, in which the 3' UTR was elucidated by SRA nucleotide BLAST of our gDNA data (the region between the two turquoise highlights) against RNA sequences of diamondback moth. The single letter amino acid codes are given above the end of the second exon (blue), with the axenome specific motif highlighted yellow. The grey region shows the 3' UTR elucidated from the BLAST and alignment of the reverse complemented RNA against our DNA sequence. The putative polyadenylation signal AATAAA is highlighted in red. Reverse primers were designed in this region to detect the likely terminus of the 3' UTR of the transcript.

4.3.9.3 DETERMINING HOW FAR THE DIAMONDBACK MOTH *B2-TUBULIN* 3' UTR

EXTENDS

Primers for five reactions were designed with the target region extending deeper into the 3' UTR (Figure 73). This facilitated an approximate method of locating where transcription stopped in the 3' UTR (Figure 74). I decided against 3' RACE as it did not matter if we overshot the 3' UTR, as when included in the transgene, transcription would behave like the endogenous gene.

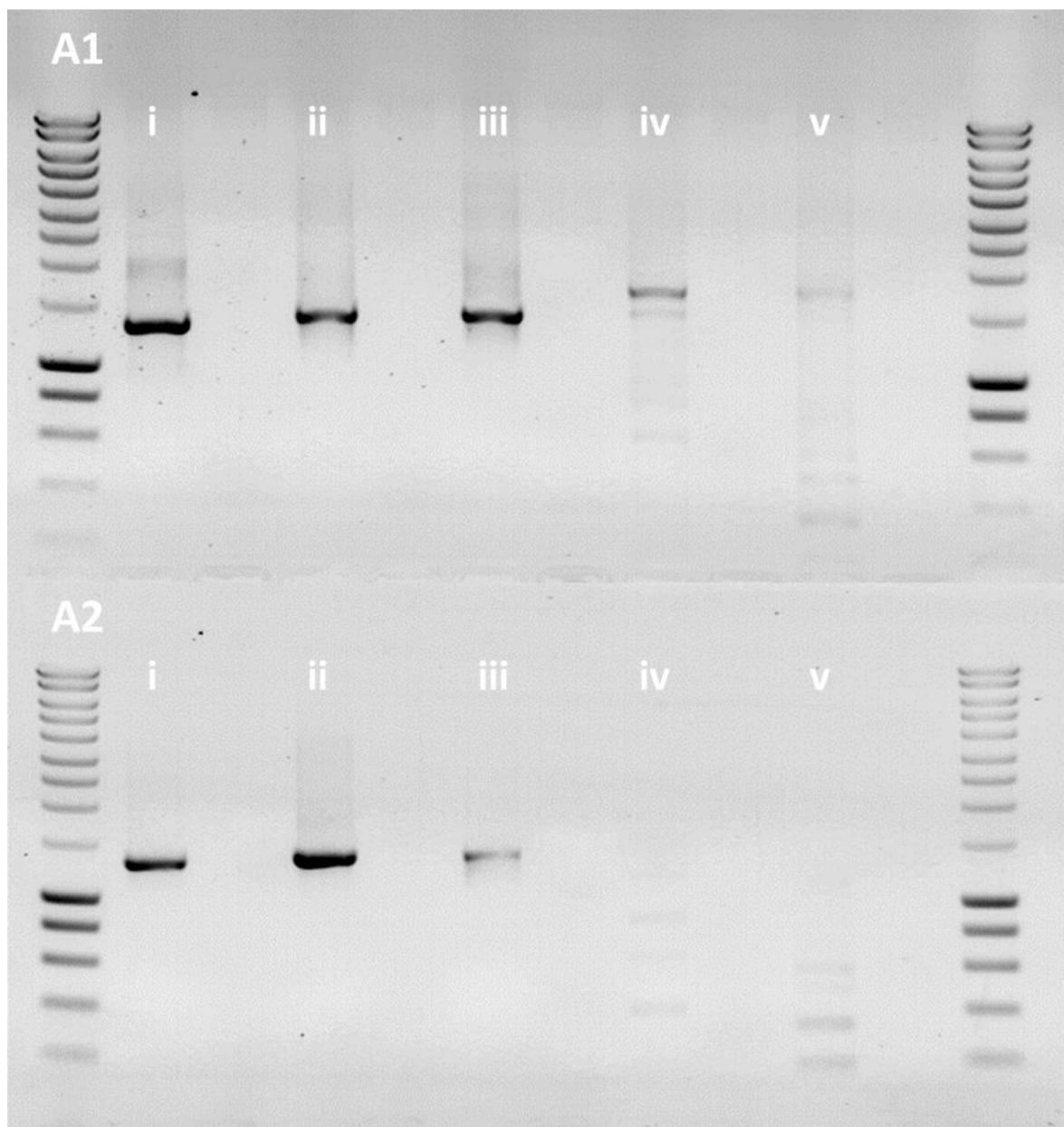


Figure 74. Where does the 3' UTR end? A series of RT-PCRs were used to elucidate how far past the stop codon the 3' UTR extends. A gel was run of 1 μ l of PCR product using 0.5 μ g RNA from OX5133A1 and OX5133A2 testes. All reactions used the same forward primer, 3976, but varied in the reverse primer used; 3977, 3978, 3979, 3980 and 3981 labelled as reactions one to five. The expected band size of the mRNA were 1295, 1369, 1394, 1442 and 1478 bp, respectively. gDNA contamination would be of an additional 189 bp in length due to inclusion of the intron. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

The RT-PCR complements the previous discussion and identified the putative 3' terminal of the mRNA as lying between the reverse primers for reactions three and four. Variation was seen between replicates with no signal in the two reactions amplifying furthest into the 3' UTR for

the second wild-type replicate. Interestingly this suggested the 3' UTR can extend beyond the first polyA signal.

A PCR was conducted with wild-type pupae gDNA using primers for reaction four targeting the 3' UTR for amplification (Figure 75). Interestingly the amplicon from one of the replicates (A in Figure 75) is larger than expected possibly due to an insertion or different allele of the gene. The PCR product from the third replicate (C in Figure 75) had the highest concentration so was cloned, purified and sequenced.

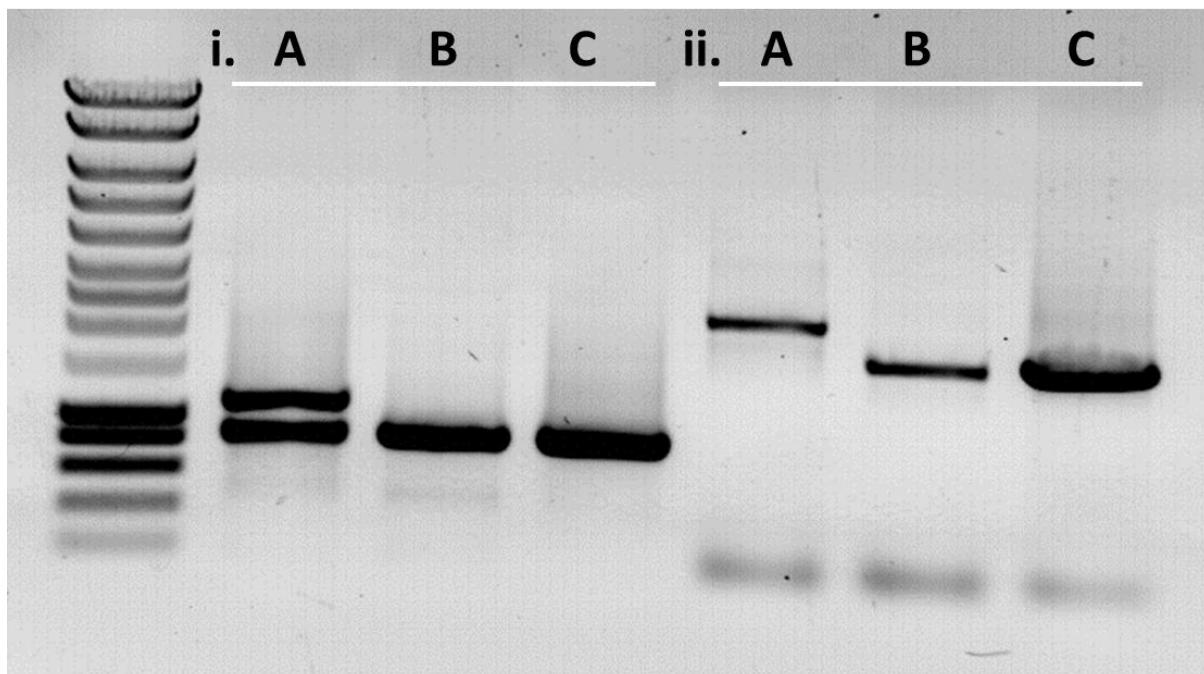


Figure 75. Amplifying the $\beta 2$ -tubulin gene including the 3' UTR in wild-type diamondback moth. The gDNA of three wild-type insects A, B and C were run in a PCR with primer pair i) $\beta 2$ -tubulin from first exon to second exon (3637 & 3638, expected size 574 bp), and pair ii) $\beta 2$ -tubulin from first exon to beyond the 3' UTR of the gene (3976 & 3981, expected size 1478 bp). Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

4.3.9.4 B2-TUBULIN 3' UTR END AND BEYOND DNA CLONING

Sequencing corroborated our practical genome Liverpool data and suggested that our cloned wild-type diamondback moth $\beta 2$ -tubulin 3' UTR was suitable for inclusion in a new sperm-marking construct design (Figure 76).

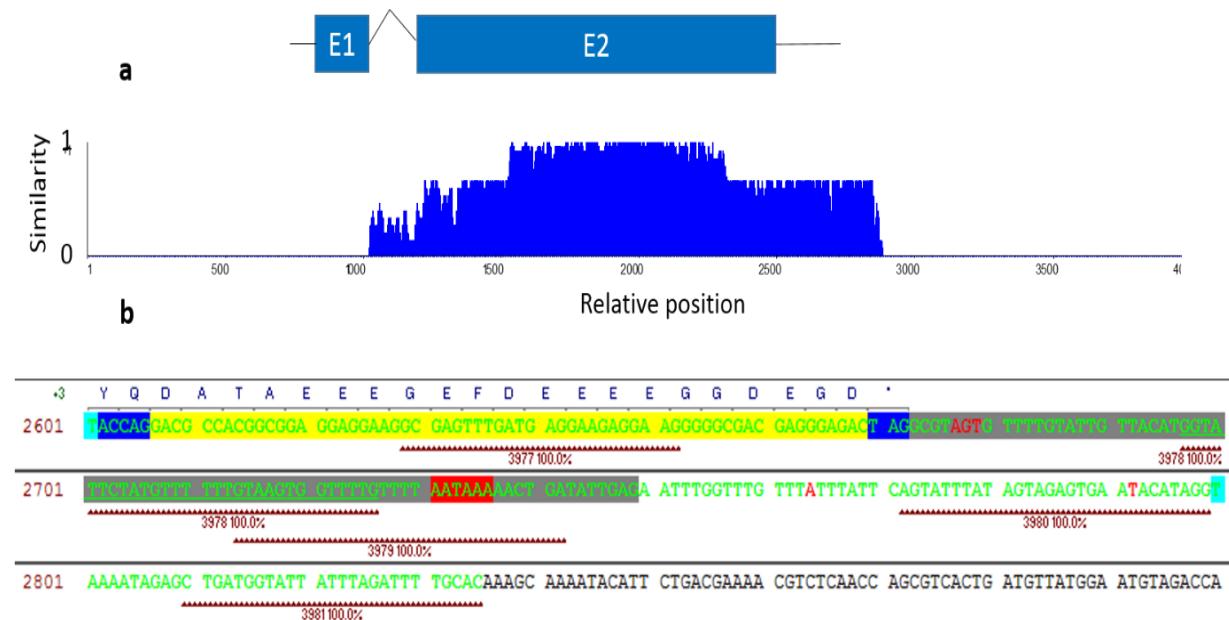


Figure 76. Comparison of genome data provided by Liverpool using shotgun sequencing with in-house amplification and sequencing at GATC-biotech. (a) The similarity of the sequences is approximately aligned and to scale with a diagram of the diamondback moth $\beta 2$ -tubulin gene to compare the similarity in the alignment with the gene features (black line, UTR; black jagged line, intron; and blue boxes, exons). The similarity is on the y-axis and uses default values where 1 means all three sequences share the same nucleotide at that position. The x-axis shows the nucleotide position or number relative to the Liverpool sequence data. (b) A close up of the sequence data of interest to help the reader (the axenome motif is in yellow with the translation above, the putative 3' UTR is in grey with discrepancies between sequence data shown in red font, the manually checked data is shown in green font going up to and including the reverse primer 3981), the red background highlights putative polyA signal. At the 3' UTR, adjacent to the stop codon, a discrepancy was detected between the Liverpool data and our sequence from the PCR where AGT was repeated in our data (7 bp after the stop codon). A deletion of a T nucleotide was found 42 bp into the 3' UTR from the stop codon in our sequence data. Two other point differences were detected in the 163 bp of the 3' UTR sequenced (plus the three unexpected AGT nucleotides).

4.3.10 FIVE TO THREE - OUR TRANSGENE IN OX5196 NOW INCLUDES MOST OF B2-TUBULIN

Rather than persisting with the all-in-one design, I opted to make use of an available reporter line OX4026C (hereafter OX4026) (see Appendix 9.1.12 for design and synthesis). Our new sperm-specific expression construct design (OX5196) included the $\beta 2$ -tubulin 3' UTR and was to

be crossed with OX4026 to provide conditional sperm-specific expression (Figure 77). Crossing heterozygotes of each line was hypothesised to produce offspring — at a 1:3 ratio — with AmCyan-sperm-expression when reared off-tetracycline.

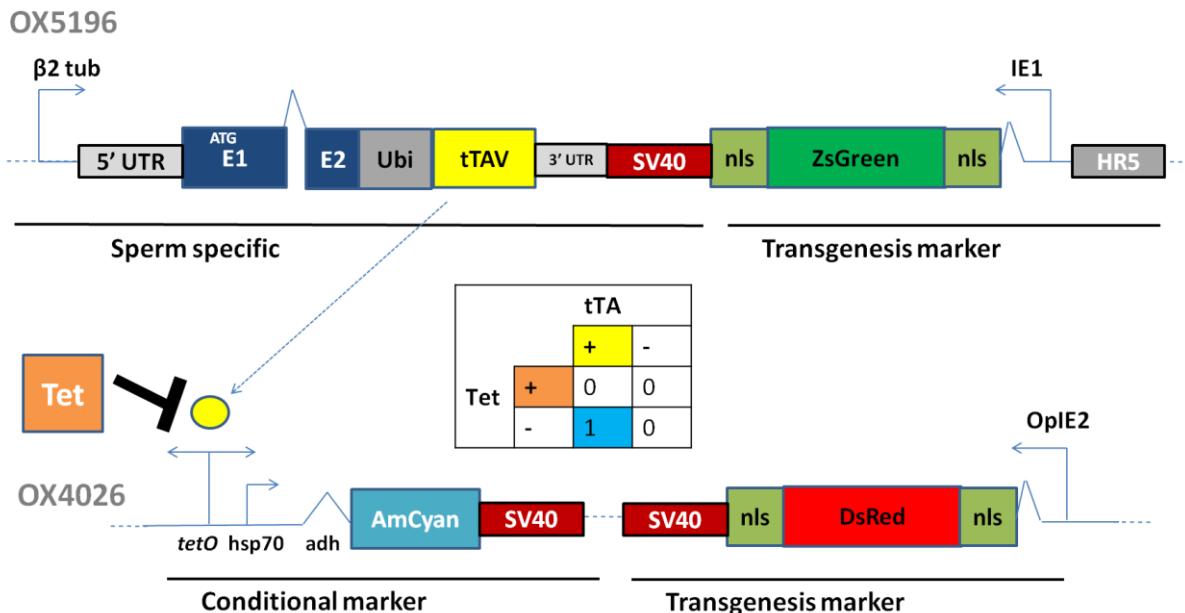


Figure 77. Crossing OX5196 and OX4026 lines will produce offspring that carry both transgenes combining conditionality and sperm-specific modules. This results in AmCyan expression in cells given the correct conditions as described by the truth table. Descriptions of each component are found in the appendix (9.1.12 OX5196). The binding of tTAV to tetO is inhibited by tetracycline and its analogues (Tet).

4.3.10.1 TRANSGENESIS

Table 24. Injection data summary for the construct OX5133 in diamondback moth. Brackets provide proportions for survival and transformation efficiency.

| Species | Construct | Embryo injections (n_1) | Injection survivors (G_0) | Independent transgenic lines (X) |
|----------------------------|-----------|-----------------------------|-------------------------------|--------------------------------------|
| <i>Plutella xylostella</i> | OX5196 | 1331 | 751 (0.56) | 5 (0.007) |

4.3.10.2 CROSSES OF PUTATIVE LINES WITH OX4026 OFF TETRACYCLINE

Heterozygote crosses off-tetracycline were set up (Methods 8.3.6.1.2) and produced genotypes inferred from the phenotypes detected by fluorescence microscopy (Figure 78). Testes were dissected and screened at the pupal stage.

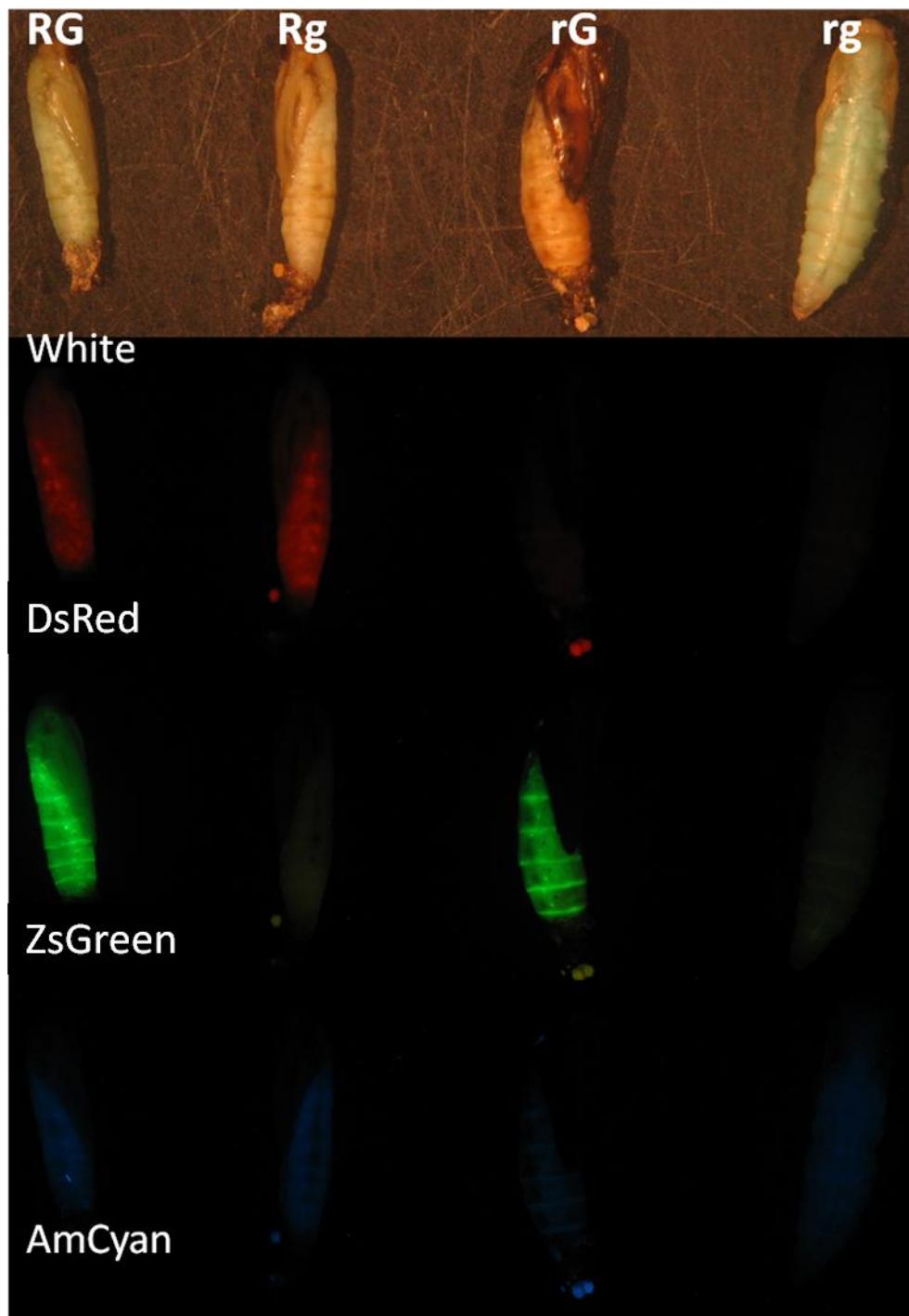


Figure 78. Crossing OX5196 with OX4026 can produce a range of phenotypes and genotypes as shown here. RG are positive for both DsRed2 and ZsGreen, Rg are positive for DsRed2, rG are positive for ZsGreen and rg are analogous to wild-type; negative for both. The bright field (white) and AmCyan lighting conditions are shown as controls. Note the RG do not appear any different for AmCyan presence.

4.3.10.3 AMCYAN EXPRESSION IN TRANSGENIC SPERM

Off-tetracycline the OX5196 transgene produced tTAV, driving expression of AmCyan to detectable levels as reported by the OX4026 transgene in two out of the three lines examined (Table 25) (two lines went extinct, lines OX5196A and OX5196D produced AmCyan marked sperm). Interestingly, some variation in brightness were observed with OX5196A appearing relatively brighter compared to the OX5196D lines.

Table 25. Offspring produced at the following ratios. Male OX5196 G₁s were used where possible to produce more offspring. The rightmost column provides the number of dissected testes examined of GR males (the denominator) and the number of those testing positive for AmCyan sperm. The number in brackets gives the number of testes examined and shown to be negative for AmCyan sperm in the respective genotypes. Some crosses produced few males, three pseudoreplicate dissections were aimed for each cross. Line OX5196B died and OX5196C cross offspring did not carry the OX5196C gene perhaps due to misclassification or mis-sexing. We were not interested in whether the expected Mendelian ratio was observed so no X² statistics are given. The expression intensity of AmCyan varied between lines with A and D producing the strongest fluorescence in that order; OX5196E derived GR were negative for AmCyan.

| ♂ | ♀ | GR | Gr | gR | gr | GR sperm Amcyan positive |
|----------|----------|-----------|-----------|-----------|-----------|-------------------------------------|
| OX4026 | OX5196A | 9 | 10 (4) | 8 (4) | 6 (4) | 5/5 |
| OX5196D1 | OX4026 | 4 | 5 (1) | 2 (1) | 4 (1) | 3/3 |
| OX5196D2 | OX4026 | 20 | 40 | 15 | 75 | 5/5 |
| OX5196D3 | OX4026 | 38 | 51 (3) | 44 (3) | 53 (3) | 6/6 |
| OX5196E1 | OX4026 | 18 | 11 (1) | 11 (1) | 19 (1) | 0/3 |

At the testes-level some somatic fluorescence from the OX5196 transgenesis marker ZsGreen was evident in G* genotype insects. DsRed2 was also detected in *R genotype but expression was not as strong. The somatic cells nuclei are positive for ZsGreen and have a branching structure that may be indicative of the tracheal system or blood supply (Figure 79).

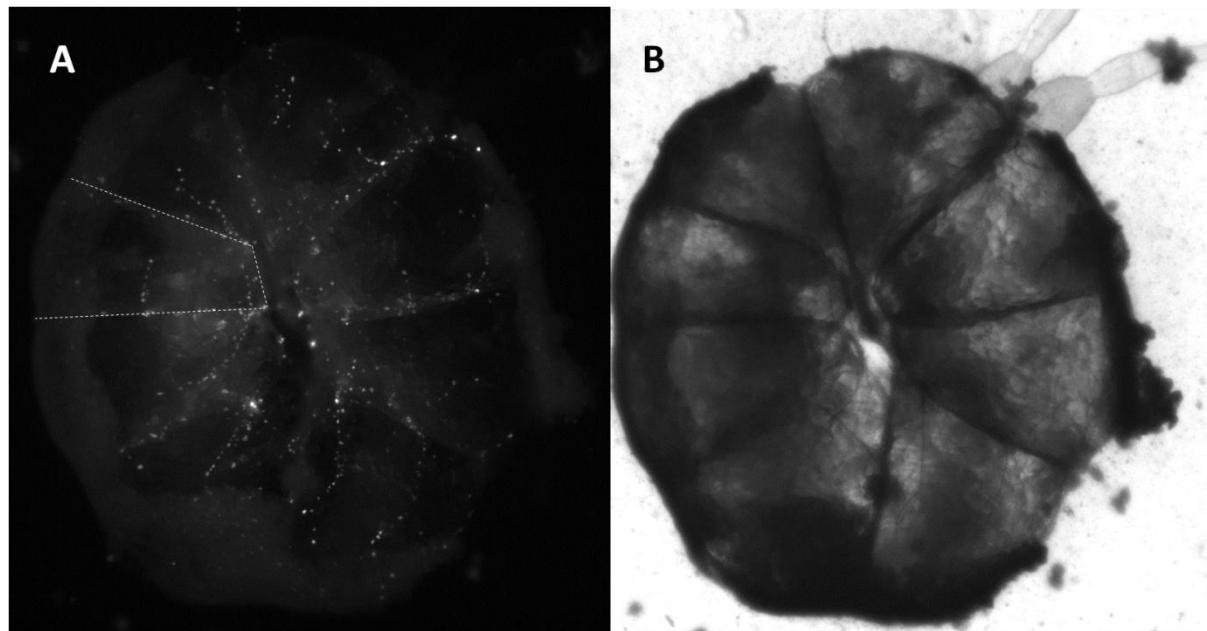


Figure 79. Evidence of somatic expression of ZsGreen on the exterior surface of the testes (x40). These testes were derived from Gr insects and provide a control for comparison with GR; both have this patterning. The testes and follicles therein must be exploded for sperm bundle access and visualisation. (A) The testes under green filter conditions and (B) the same testes under white light. Note the eight follicles (one follicle has a dashed line around it) and vas deferens in the top right of the picture. The diameter of the testes is approximately 900 μm .

Exploding the testes is required to assess $\beta 2$ -tubulin-AmCyan expression. Sperm bundle AmCyan fluoresce was only detected in GR insects and was distinguishable from the somatic expression (Figure 80A-B). Magnifying the sperm bundles further we compared GR sperm with Gr sperm to determine if ZsGreen or DsRed2 from the transgenesis markers were affecting our conclusions (Figure 80C-F). Gr insects' sperm were used as a negative control and were negative for fluorescence under AmCyan filter conditions (Figure 81). The other genotypes; gR and gr were also negative for sperm fluorescence.

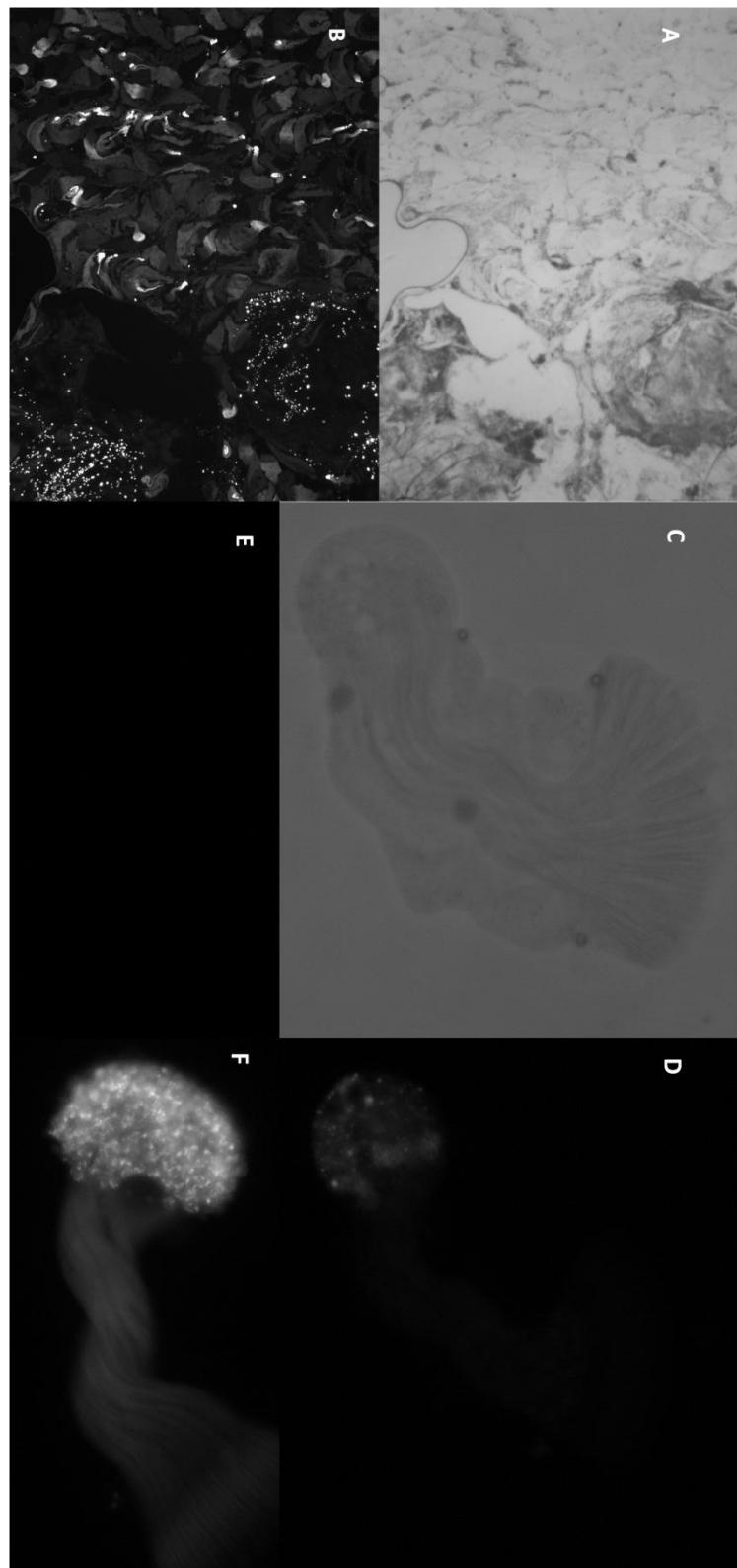


Figure 80. β 2-tubulin can drive sperm specific expression. AmCyan positive sperm bundle fluorescence compared to Hr5-*Ie1*-ZsGreen fluorescence in an exploded testes of a GR insect. (A) Bright field x100 image of OX5196A x OX4026 GR exploded testes (B) and blue light / green filter. Both AmCyan and ZsGreen are detectable under the same lighting conditions but can be distinguished by the nls-ZsGreen causing pin-prick like foci (dashed line) of light compared to the sperm head fluorescence (solid arrow). The ZsGreen appears to cluster around the visible branches in the bright field image; possibly blood supply or tracheal system (dashed line). Insects were

reared off-tetracycline. (C) Bright field x400 (D) AmCyan positive line OX5196D crossed OX4026 GR, moderate but detectable expression levels under AmCyan screen. (E) DsRed2 lighting conditions of a the sperm bundle in F, no bleed through from AmCyan or ZsGreen detection (F) AmCyan of high expression line OX5196A crossed OX4026 GR. Within each line expression seemed consistent with OX5196A always bright and OX5196D more moderate expression levels. The AmCyan is translated in the head region near the nuclei and does not appear to be transported into the tail. Note: A & B are paired, C & D are paired. GR sperm where the only genotype to display AmCyan expression off tetracycline (Gr, gR and gr did not).

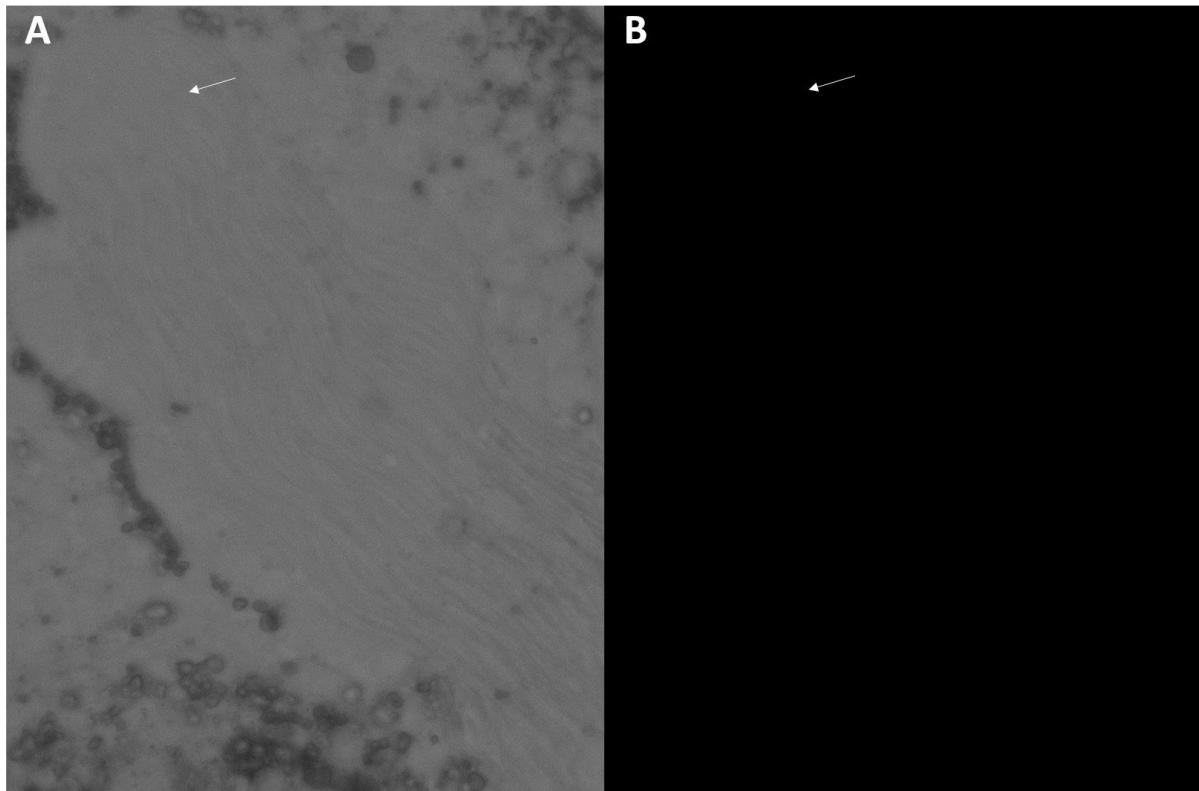


Figure 81. *Hr5-*Ie1*-ZsGreen*, the transgenesis marker, is not detectable in Gr sperm bundles. This suggests the AmCyan signal is real and unique to GR insects. OX5196 produces tTAV in the sperm which regulates AmCyan expression via *tetO*. (A) Bright field x400 (B) under blue light and green filter conditions.

Interestingly this transgenic system provides an insight as to the timing of endogenous translation of stored $\beta 2$ -tubulin transcripts during spermatogenesis in diamondback moth. Translation begins after meiosis and during spermiogenesis, when the sperm have differentiated and taken on the distinctive elongated shape (Figure 82). This appears late on during the elongation stage prior to maturation. This is plausible given the role of $\beta 2$ -tubulin in the motile axenome. This is comparable to the timing of transcription and translation of Medfly $\beta 2$ -tubulin described by Jin (2011).

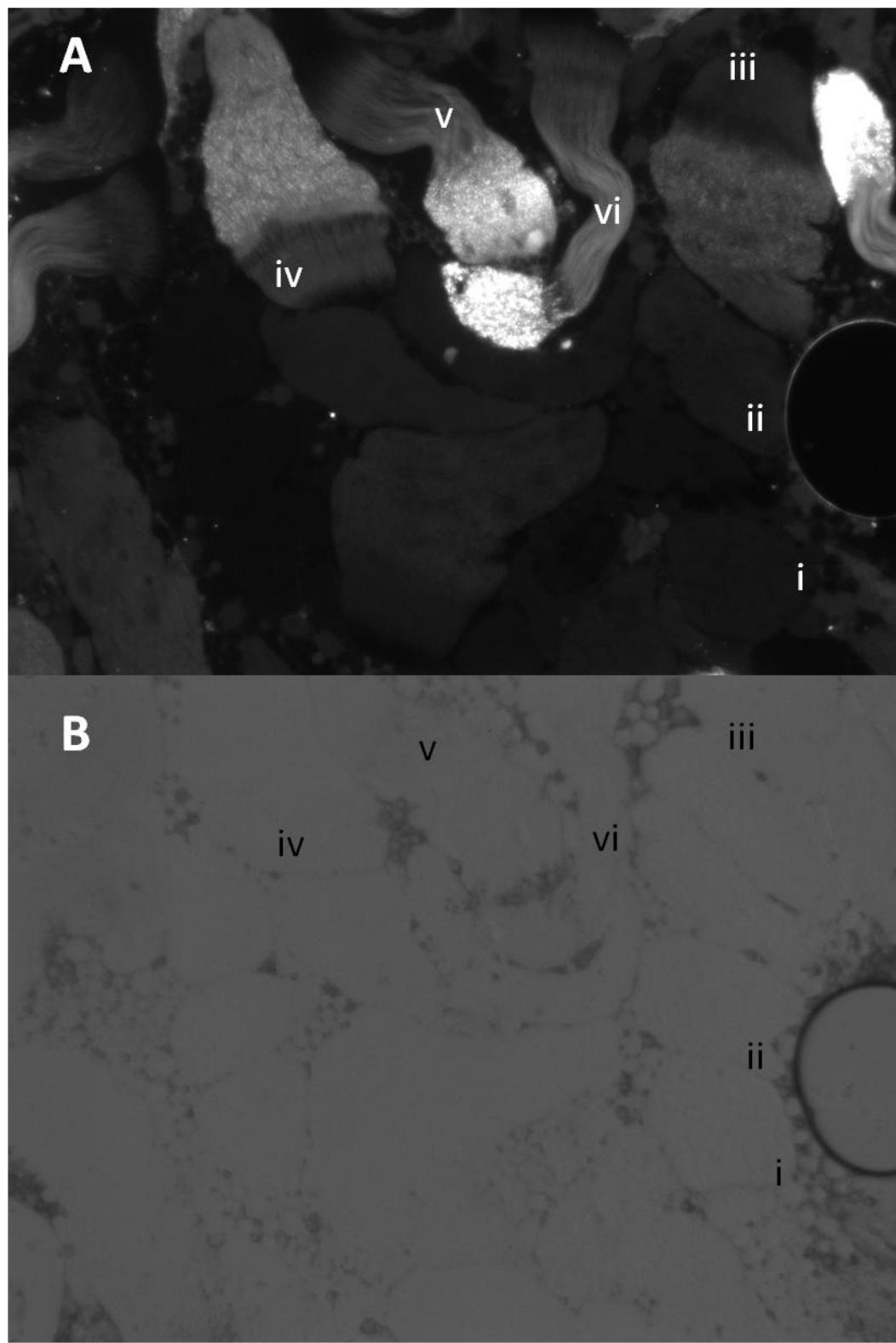


Figure 82. Tracking $\beta 2$ tubulin-AmCyan translation to infer endogenous gene translation timing in spermiogenesis. Different stages of spermatid maturity can be seen within the cysts labelled here (based on sperm development); we see development from circular spermatid to elongated spermatids while bundled and contained in their respective cysts through i to vi. The earlier stages (circular cysts) show no AmCyan translation implying $\beta 2$ tubulin transcripts are yet to be translated. As the spermatid nuclei within each cyst migrate during spermiogenesis they also translate the transcripts into the respective AmCyan as flagella require the endogenous $\beta 2$ tubulin to produce

mature sperm. $\beta 2$ -tubulin extends the microtubule structure from the base, this may be evident in vi where we see some AmCyan putatively incorporated into the base around the start of the tail region. Sperm bundle iv shows the nuclei as being needle shaped and longer than the flagellum which has just started to assemble, after this stage, in v, the nuclei appear to take up proportionally less of the sperm at the expense of the growth of the tail. Individualisation of the sperm may occur between stages ii and iii.

Prior to spermiogenesis the nuclei migrate to one pole of the cyst (in *Drosophila*, Reimer, 2012).

The $\beta 2$ -tubulin-AmCyan translation continues to illuminate this maturation and cell restructuring through different stages. As the cysts containing the maturing sperm change shape from circular to an irregular ellipsoid we see the spermatids maturing and becoming distinct and individualised. Unsurprisingly, the translation is concomitant with the growth of the flagella probably due to the role of $\beta 2$ -tubulin protein in motile axoneme construction.

4.3.10.11 TRANSLATIONAL CONTROL BY THE 3' UTR: THE ENDS SPECIFY THE MEANS

Arguably the more profound implication of this discovery is the role of the endogenous 3' UTR regulating the translation of the diamondback moth $\beta 2$ -tubulin. This Chapter has demonstrated that the non-coding DNA provides transcription specificity to the sperm prior to meiotic arrest and the 3' UTR regulates translation at spermiogenesis. The transcripts are not required immediately (White-Cooper & Davidson, 2011) so persist in the spermatids until the elongation and maturation stage where a putative signal binds to the 3' UTR and facilitates translation. Perhaps there is a master regulatory mechanism with this hypothesised translation factor activating a cascade of transcripts involved in the complex processes of spermiogenesis? Hitherto this finding, this temporal-uncoupling between transcription and translation was assumed to be achieved by repression of the transcripts at the 5' or 3' end (Steger, 2001).

Most sperm-specific genes are assumed to be transcribed and held in a translationally repressed state, for several days, until required. This was demonstrated in *Drosophila* $\beta 2$ -tubulin (β Tub85D), where a fragment consisting of only 53 bp of promoter region, plus the first 71 bp of the 5' UTR was sufficient to confer testis-specific expression on reporter genes. Here we show that the 3' UTR is necessary for the translation of the transcript. This contrasts to the Diptera where the 3' UTR of the species $\beta 2$ -tubulin is not required for sperm marking (Scolari *et al.*,

2008; Zimowska *et al.*, 2009; Jin, 2011). Again diamondback moth surprises us, with not only a difference to the Diptera for the transcription of the gene but for the translation as well.

Interestingly the 3' UTR control element required for diamondback moth $\beta 2$ -tubulin is small at less than 77 bp. This is typical for testis-specific control elements; perhaps it is important they are small thus facilitating new gene duplicates to be expressed in testes (White-Cooper, 2010).

Speculation can be made about the type of control occurring. It is unlikely that micro RNAs are regulating the translation as typically repression is achieved by forming imperfect complementarity with elements in the 3' UTR of their targets (Kuersten & Goodwin, 2003). Here translation is only realised with the addition of the 3' UTR, suggestive of a protein-mediated interaction (or the failure of the SV40 polyA signal). Perhaps a protein-complex binds to the 3' UTR and facilitates the formation of an RNA loop between the transcript termini (Maumbder *et al.*, 2003). It is hypothesised that this protein would be expressed during spermiogenesis and may be involved in the 3' UTR translation regulation of a suite of transcripts required by the cell during this complex differentiation from spermatid to mature sperm. This provides us with two plausible models of diamondback moth $\beta 2$ -tubulin translation regulation (Figure 83).

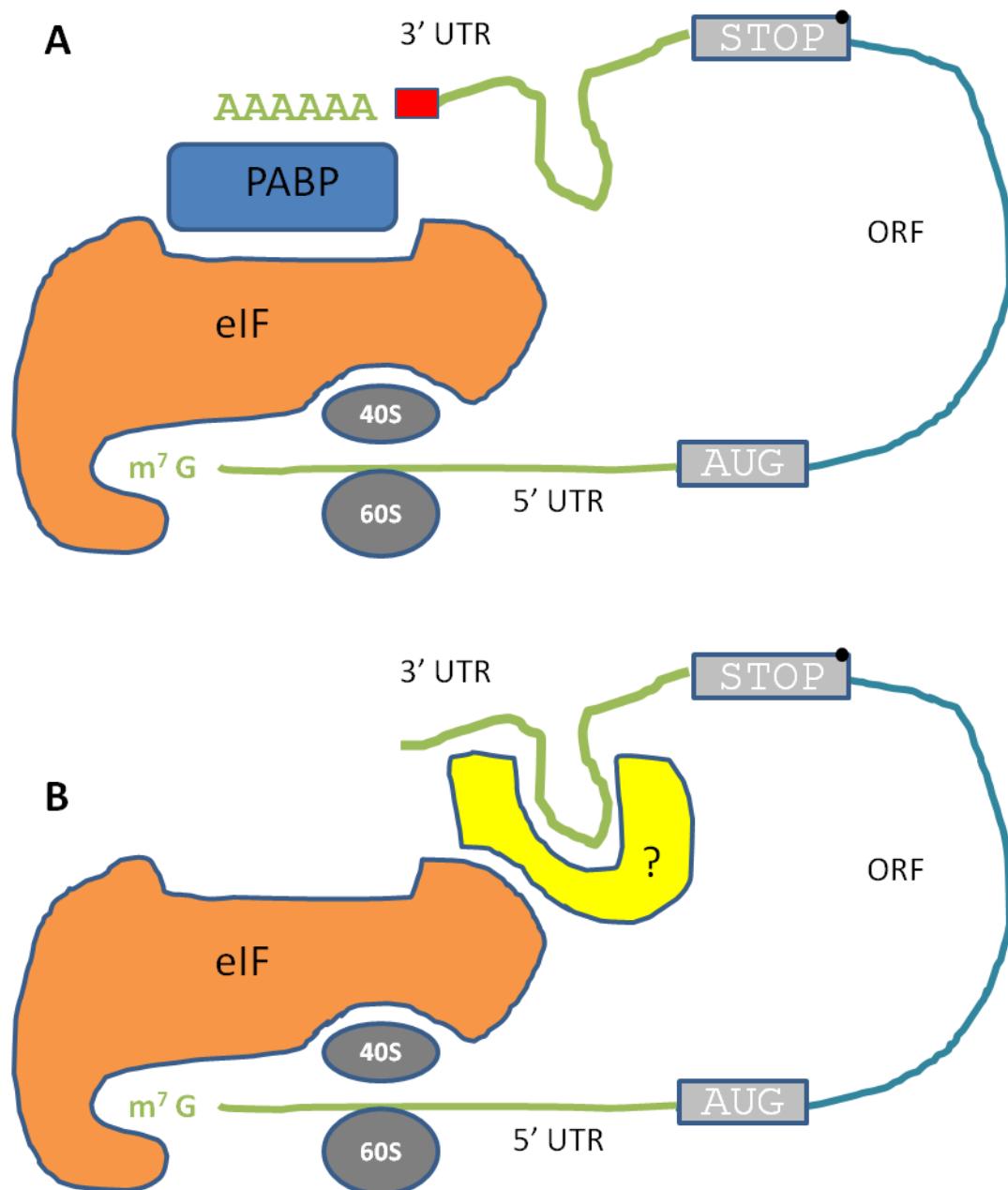


Figure 83. Models of putative translational control of diamondback moth $\beta 2$ -tubulin involving transcript circularisation. 'Circular' or 'closed-loop' mRNA model showing two plausible models given the necessity of the endogenous 3' UTR for translation. Circularisation is required for translation initiation complex formation and ribosome attachment and scanning. A) Circularization mediated by poly(A)- binding protein (PABP) binding to both the poly(A) tail and eukaryotic initiation factor (eIF) of the initiation complex. In this version of the model our earlier constructs failed as SV40 does not contain an appropriate poly(A) signal recognised by the cell during spermatogenesis. The addition of endogenous 3' UTR provided correct poly(A) signal and subsequent translation. This version seems less likely as RT-PCR showed high transcription levels which seem unlikely given no poly(A) tail would negatively affect the stability of the transcript. Red box – poly (A) signal B) The poly(A) tail ensures stability but does not affect circularisation, this is provided by the endogenous 3' UTR recognising the hypothesised yellow protein (labelled ?). Both termini are involved in recognition at the 5' (m^7GpppN cap) and 3' end. Recognition and binding to respective proteins at both termini causes a conformational shape change and ribosome attachment (40S and 60S subunits combine to form 80S ribosome). ORF - open reading frame. (After Mazumbder *et al.*, 2003). Sequencing the 3' end of transcripts of earlier SV40 only containing transcripts for the poly(A) tail could test model A. Alternative models may exist that explain the data.

4.4 CONCLUSIONS

This chapter started by introducing a breakdown of the chapter objective into manageable hypotheses to be assessed in turn.

- The orthologue for $\beta 2$ -tubulin can be found in diamondback moth and confirmed by RT-PCR and *in situ* hybridisation of the testes to show testis specificity.
- $\beta 2$ -tubulin promoter region in diamondback moth can be identified and provide sperm specific transcription.
- A fluorescent protein under control of the $\beta 2$ -tubulin promoter will result in fluorescent sperm and/or testes post-late-larval stage.
- A suitable protamine like gene can be identified in diamondback moth for use in the paternal effect system.
- A “paternal effect system” will provide effective sterility phenotype in male diamondback moth using the *Foq1* nuclease.

The chapter will finish by reflecting on these hypotheses and synthesising evidence from this chapter and the literature to provide conclusions as to their status.

Through the use of RNA-Seq analysis of different tissue types — confirmed by RT-PCR — candidate genes for use in a paternal effect system were short-listed. Five genes with high testis-specific expression were chosen for analysis with RISH of the testes to determine temporal-spatial transcript profiles in diamondback moth pupae. The method was successful in three of the five genes demonstrating the versatility of the method and its application in a novel species, developed originally in *Drosophila* (Morris *et al.*, 2009).

$\beta 2$ -tubulin has proven to be a model gene of choice when investigating sperm specific expression in the testes of insects during spermatogenesis. This gene was confirmed to share similar testes-specific transcription profiles in diamondback moth as in other insects. Transcript was still present at the pupal stage suggesting that translation continues into adulthood. Upon

translation these transcripts would become destabilised, and so the time at which a transcript disappears during spermiogenesis correlates with the time that the protein is produced (White-Cooper, 2010). Development of the OX5196 and OX4026 crosses reared off-tetracycline facilitate comparison between $\beta 2$ -*tubulin* transcription (RISH-DIG probes) and translation (AmCyan fluorescence) in diamondback moth. The amount of signal by follicle position could be compared with an inverse signal expected on comparison.

Due to the superficial similarities with other species in which a paternal effect system has been developed at Oxitec I decided to build a similar construct for use in diamondback moth. Accordingly the endogenous $\beta 2$ -*tubulin* promoter region was cloned and used to drive the desired testes-specific expression. The integrity of the sequence was assessed using 5' RACE. Otherwise the construct structure was duplicated from successful Medfly and mosquito construct designs.

Through trial and error — using insect transgenesis — it was determined that the promoter region alone (~4.5 kbp) and 5' UTR (shown to be intron free) were not sufficient to drive testes-specific transcription (Figure 84). This is a notable contrast to the Diptera where typically a small promoter region and part of the 5' UTR is usually sufficient to drive transcription in male germline genes (White-Cooper, 2010). This suggests that although the $\beta 2$ -*tubulin* coding region is highly conserved the regulation of the gene may be less so.

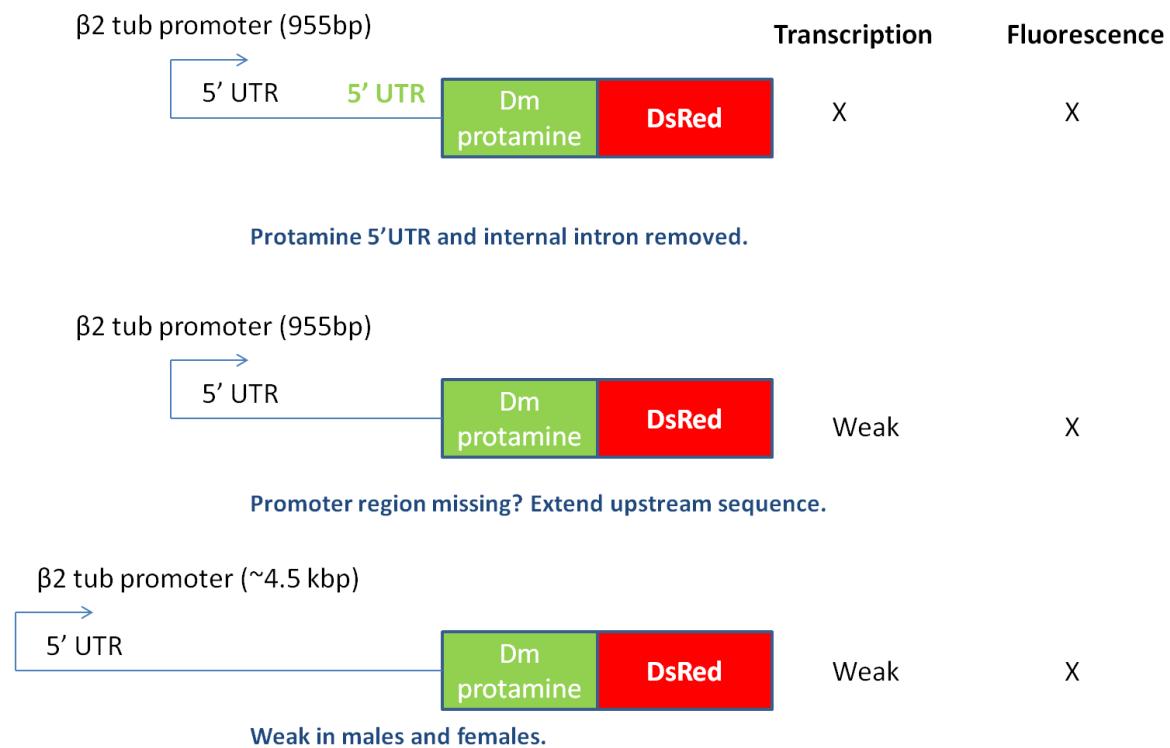


Figure 84. A research dead end – these constructs were designed on the assumption that what worked in the Diptera could be transferred into diamondback moth. However, the many components approach resulted in not much being learnt between constructs. The constructs are presented chronologically from top to bottom; OX4703, OX4804, OX5067. The blue text describes the outcomes and actions to design the next construct. The protamine is derived from *Drosophila melanogaster*. Not only was transcription weak but it was non-specific with transcription in non-testes tissues.

This was confirmed by the re-design of the construct to contain additional features of the $\beta 2$ -tubulin gene. The first exon, intron and part of the second exon were included upstream of the reporter protein. It was suspected that an intronic enhancer may provide the testis-specific transcription in diamondback moth, as shown in another lepidopteran, *Heliothis virescens* (Raff *et al.*, 1997). This approach was vindicated with OX5116 producing analogous levels of the transgene transcript as the endogenous gene (Figure 85). This suggests a functional role in the intron of the gene, particularly when considering that alternative splicing is not known in the tubulins (Nielsen, 2010). This contrasts with the utility of introns in dipteran germline genes, typically they are considered detrimental to transcription (pers. comm., Helen White-Cooper). This region may also contain a repressor region for somatic cell regulation, as non-testes transcription was observed when it was absent from the transgene.

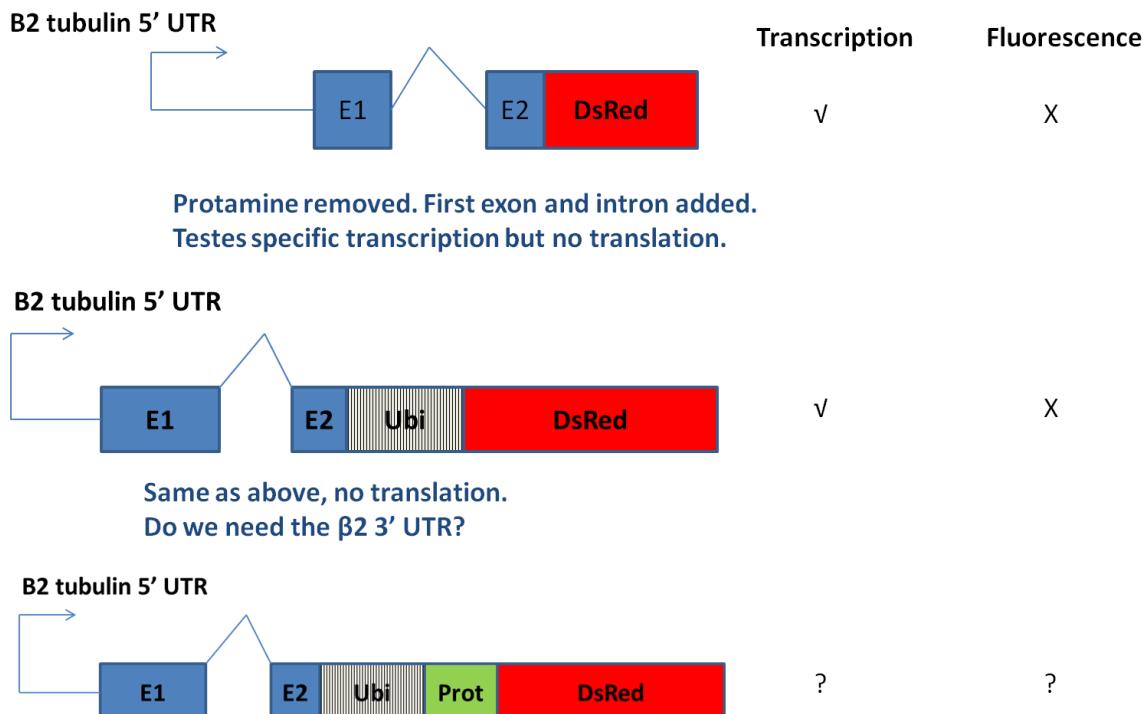


Figure 85. Testes specific expression was achieved with the above design. A region between exon 1 and the start of exon 2 were necessary for transcription. Correctly spliced and sequenced transcript was produced but no DsRed2 protein was detected. The constructs are presented chronologically from top to bottom; OX5116, OX5133, OX5135. The promoter region was the full 4.5 kb region of the $\beta 2$ -tubulin gene plus the 5' UTR. The protamine in OX5135 is *Drosophila melanogaster* derived.

What is the possible cause of such a difference between the Diptera and diamondback moth? I suspect it is due to the evolutionary constrained nature of the gene in the different species. In the Diptera specifically the *Drosophila* clade, the evolutionary stasis is due to functional constraint: stringency in the structure / function relationship between $\beta 2$ -tubulin and the sperm tail axoneme does not permit viable variation in the protein. Synergism as well as possible constraints resulting from having to support such a long axoneme, may act to keep the entire protein from evolving.

$\beta 2$ -tubulin has evolved in dipterans since mosquitoes and fruit flies shared an ancestor, though more slowly than in lepidopterans. This rate difference could be explained by the absence and presence, respectively, of a testis-specific α -tubulin (Figure 86).

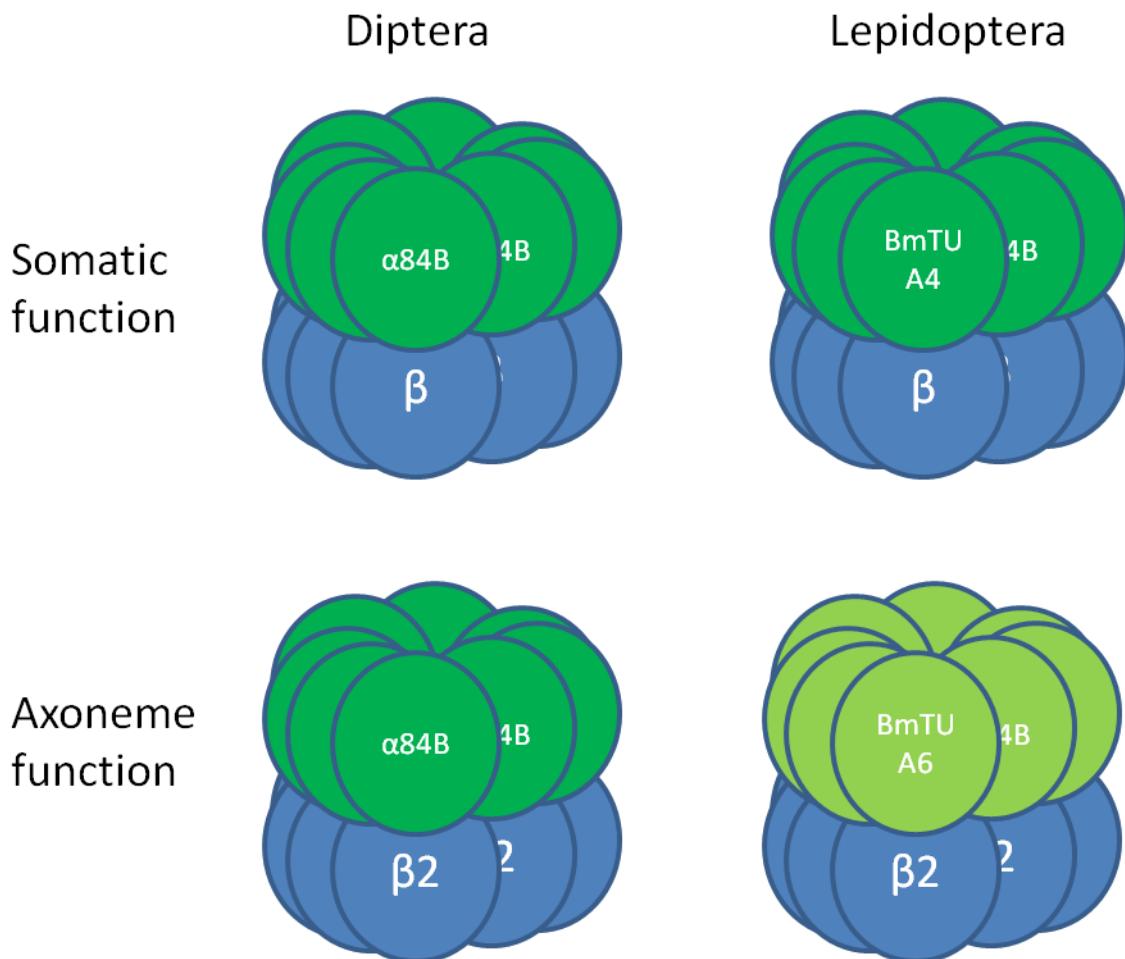


Figure 86. *Drosophila* use a major α -isoform ($\alpha 84B$) to support both somatic and axoneme microtubule function, whereas *Bombyx mori* divide these functions between a major isoform ($BmTUA4$) that supports somatic function, and a testis-specific α -isoform ($BmTUA6$) that supports the axoneme. The major α -isoforms $BmTUA4$ and $Dma84B$ are 99% identical; however, $BmTUA6$ and $Dma84B$, both of which support motile axonemes, are only 78% identical. If use of a major α -isoform and testis-specific β -isoform in the axoneme is conserved among Dipterans, then Dipteran testis-specific β -tubulin has evolved with essentially the same α -tubulin for 265Myr (Nielsen et al., 2006).

Any mutation in the $\beta 2$ -tubulin will affect its compatibility with its partner α -tubulin. In the Diptera this is particularly constrained and makes the chances of a beneficial mutation less likely as a change would have to be compatible with both α -tubulin isoforms due to its functional role in the soma and axoneme. In contrast the $\beta 2$ -tubulin in Lepidoptera need only synergise with its axoneme specific partner, rendering heritable variation more likely. This division of labour possibly allows the axoneme to explore more functional space, to be more evolvable, than possible when one member of the dimer has multiple responsibilities.

This constrained evolution reveals why $\beta 2$ -tubulin may not be a “Biobrick” of the paternal effect system. If it is more evolvable in other insect orders then its regulation is also likely to be more evolvable and different compared to the consistency seen in the Drosophila clade. This is probably why the use of this gene in diamondback moth proved relatively difficult. The introns were not anticipated to regulate the testis-specific transcription and surprised many researchers in this field.

Despite producing the desired transcription of the transgene in a testis-specific manner with high transcript levels, no fluorescence was detected. Following our best efforts to remedy this using the Ubiquitin system (to solve putative DsRed2 folding problems) as well as trying to measure the protein directly with immuno-blot techniques, experiments were inconclusive.

Most transcripts in *D. melanogaster* that are required to produce proteins for sperm tail are synthesised in primary spermatocyte nuclei, and stored in a translationally repressed state for several days, until late in spermiogenesis (White-Cooper, 2010). I suspected that the 3' UTR may serve a role in translation regulation in diamondback moth $\beta 2$ -tubulin in contrast to the Diptera model. Accordingly the endogenous 3' UTR was cloned, sequenced and added to our construct. The hope being that the 3' UTR activates translation compared to the SV40 failure used in earlier constructs. This approach was successful with the endogenous 3' UTR necessary for translation of the transgene providing sperm-marking. The 3' UTR is hypothesised to provide a critical role in RNA circularisation necessary for closed-loop transcript formation necessary for recruitment of the translation initiation complex and translation machinery.

Ultimately we achieved testis-specific transcription and translation of a transgene in diamondback moth. We did not get the opportunity to test the effect of a *Dm protamine* (diamondback moth protamine could not be identified) on the nuclear localisation of a protein nor the efficacy of *FoqI* in introducing sterility into males.

CHAPTER 5 – THE DEVELOPMENT OF A REPRESSIBLE GENETIC SEXING MECHANISM IN *TRIBOLIUM* *CASTANEUM*



5.1 BACKGROUND

5.1.1 SUITABILITY FOR SIT

The Coleoptera are the most speciose of the insect orders giving rise to many economically relevant pest species. The Coleoptera are a neglected insect order for SIT research (Suckling et

al., 2013). Genetic control holds potential as another tool in humanity's arsenal against these pests. Before attempting to apply this technology to a specific pest problem it is prudent to test it on the model coleopteran, the red flour beetle, *Tribolium castaneum* (hereafter, *Tribolium*). This approach has proven fruitful in the Diptera with advances and insights with the genetic control of *D. melanogaster* being transferred to other pest species thereafter (Handler *et al.*, 1998).

Despite its model organism status, it is unlikely that *Tribolium* will be amenable to SIT as it fails to meet some of the criteria laid out by Knipling (1955). Of greatest concern is the polyandrous nature of females; an average of 3.2 matings per female was estimated by Pai & Bernasconi (2008). Furthermore, both sexes feed and are destructive to stored products by direct feeding or reducing product quality.

The purpose of developing transgenesis within this species is to demonstrate the principle for the genetic control and sexing of coleopteran pests, which include the boll weevil (*Anthonomus grandis*), the red palm weevil (*Rhynchophorus ferrugineus*), the rice water weevil (*Lissorhoptrus oryzophilus*), the Colorado potato beetle (*Leptinotarsa decemlineata*), the coconut hispine beetle (*Brontispa longissima*) and many other economically damaging pest species affecting global food security.

Current methods of control of *Tribolium* rely on fumigation strategies suited to the enclosed flour and rice mills where the pest causes losses (Buckman *et al.*, 2013). The recent evolution of resistance to insecticides, as well as the phasing out of some of the more effective fumigants, may render these control methods unsustainable. Exploring genetic control will provide a sustainable alternative as part of an IPM approach.

5.1.2 SUITABILITY FOR GENETIC CONTROL AND A SEXING STRAIN USING RIDL

A common challenge encountered when engineering RIDL strains in a new insect species is the lack of characterisation of gene expression systems capable of conferring female-specific expression at early developmental stages, hindering the development of sexing strains. The availability of the sequenced *Tribolium* genome provides a useful starting point for transgene design; with *Tribolium* considered the “third best invertebrate for genetic and molecular studies” (Denell, 2008). Genes involved in the sex determination pathway serve as an obvious starting point and may offer genes which are alternatively spliced in the sexes.

Sex determination and development regulatory genes are surprisingly non-conserved among closely related model organisms when compared to other key developmental and regulatory genes (Suzuki, 2010). In the insects, sex determination is achieved through two alternative cascades of gene regulation; the master switch controlling each cascade varies from order to order (Sanchez, 2008).

5.1.3 SEX DETERMINATION

5.1.3.1 IN INSECTS

In general there are three primary genes involved in insect sex determination (Sanchez, 2008): *sex lethal (sxl)* (Traut *et al.*, 2006), *transformer (tra)* (Shukla & Palli, 2012) and *doublesex (dsx)*. Comparative studies in insects have found that alternative splicing reiteratively evolves to control expression of the key sex-determining genes (Arunkumar & Nagaraju, 2011; Salz, 2011).

Exploiting sex-alternate splicing offers the potential for a genetic sexing method as demonstrated in other species (Papathanos *et al.*, 2009) and could be achieved by using regulatory sequences from one of a number of these sex determination genes combined with

the RIDL system (Gong *et al.*, 2005; Jin *et al.*, 2013), with tTA (or tTAV) acting as both transactivator and lethal effector (Gallia & Khalili, 1998).

5.1.3.2 SEX DETERMINATION IN *TRIBOLIUM*

A hierarchy of genes controls somatic sexual development in *Tribolium* with alternative splicing of genes being the genetic regulatory method of choice to generate the two sexes. The default sex in *Tribolium* is male, in that knock-down of some of the key genes involved in the sex determination cascade can produce all-male progeny (Shukla & Pali, 2012).

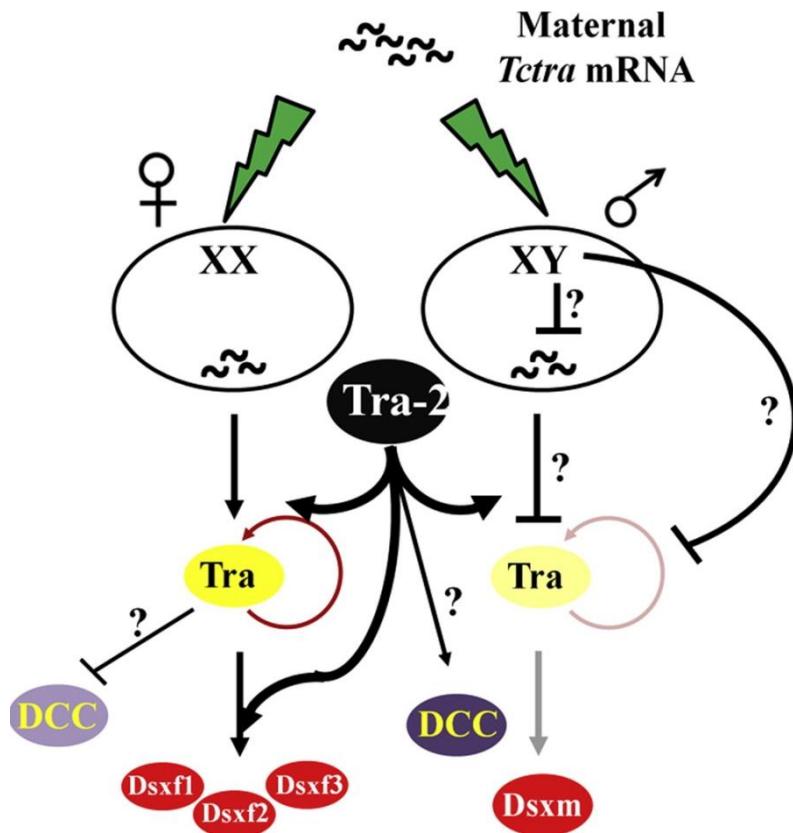


Figure 87. Proposed model of the sex-determination cascade in *Tribolium castaneum*. Transformer2 (*Tctra-2*) is required for the female-specific splicing of *transformer* (*tra*) and *doublesex* (*dsx*) pre-mRNAs in females. *Tctra* is also required for the male-specific splicing of *Tctra* pre-mRNA in males. Both *Tra* and *Tra-2* may be required to inhibit the formation of dosage compensation complex (DCC) in females whereas in males *Tra-2* alone is required for the formation of dosage compensation complex. This is based on the fact that *Tctra* knockdown leads to death of only females (Shukla and Pali, 2012 and this chapter) whereas *Tctra-2* knockdown leads to death of both female and male. After Shukla and Pali (2013).

doublesex (*dsx*) is the last gene in the sex determination cascade. Male and females produce distinct splice variants of the gene (Suzuki, 2010). This differential expression permits a means

of separating the sexes by inserting an effector in a sex-specific exon of a *doublesex* minigene so that, during the sex determination process, the effector is expressed (Fu *et al.*, 2007). The differential splicing of the *doublesex* pre-mRNA is modulated by functional Tra protein (derived from the splicing of the *transformer* gene (*tra*) transcript) and an association with Tra2 (Shukla & Pali, 2012).

5.1.4 PROMOTER CHOICE

The choice of promoter driving marker expression in *Tribolium* has proven pivotal to transformation success in this species, with several research groups concluding that high-level transgene expression requires endogenous regulatory sequences (Siebert *et al.*, 2008; Schinko *et al.*, 2010). *Tribolium* α -tubulin-1 promoter has been used reliably to promote marker expression following *piggyBac*-facilitated insertion. However, when developing transgenic technologies in a new species it is advantageous to develop an arsenal of promoters that work to allow flexibility in transgene and marker expression especially when a combination of transgenes are included in the construct. This should be considered throughout the project with the testing of promoters that have a history of application to diverse insect orders, such as the *Opie2* promoter (Theilmann & Stewart, 1992; Pfeiffer *et al.*, 2010).

5.2. OBJECTIVES

Genetic manipulation of *Tribolium* has been previously undertaken by other groups. The use of published genome information should be used to generate transgenic strains with a genetic-sexing module.

In chapter 1 I introduced my objectives for *Tribolium*:

- To engineer a transgenic strain of *Tribolium* using a fluorescent protein marker linked to an endogenous promoter, reproducing research from the literature.
- To develop a conditional female-specific lethal genetic system in *Tribolium*.
- To engineer repressible male sterility phenotype in *Tribolium* ("paternal effect system").

Compared to diamondback moth, where significant progress had been made towards third generation technology, for *Tribolium* we must start from the beginning. Prior to the sperm lethal system we had to develop the precursor modules including proof-of-principle of genetic sexing and conditionality.

These overarching objectives can be broken down into smaller research goals and hypotheses:

- Achieve germline transgenesis of *Tribolium* using *piggyBac*.
- Determine if the tTAV-*tetO* binary system is viable in *Tribolium*; is the system tetracycline-repressible through the diet?
- Identify a suitable endogenous sex determination gene and validate that it is alternatively spliced for female-specific expression.
- Generation of a repressible female-specific lethal transgenic line, based on suitable alternative splicing of a sex determination gene.
- Identify suitable promoter for sperm-specific transcription.
- Generation of a sperm-lethal transgenic line by combining these modules.

5.3 RESULTS AND DISCUSSION

The work described in this chapter aimed to design and implement the three genetic modules that are pre-requisite for the paternal-effect system (Figure 88).

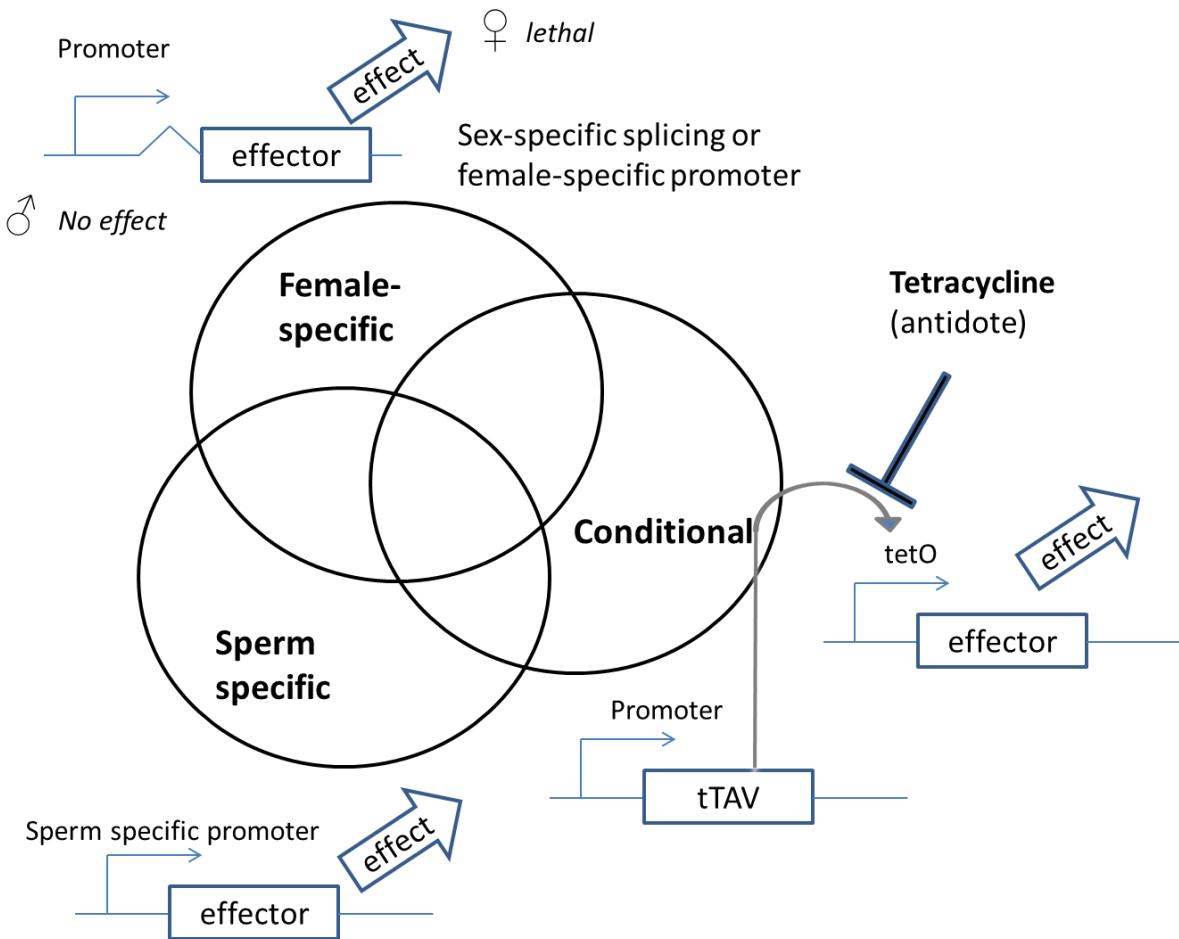


Figure 88. The paternal-effect system using synthetic biology: combinatorial control of gene expression with defined modular control elements. In the constructs described here, the sperm specific promoter, sex-specific intron (wavy line) and tetracycline-repressible trans-activation mechanism (tTAV or alternatives bind and activate the tetO enhancer in the absence of tetracycline) act as independent control elements; logic gates which combine specific inputs (respectively Tissue and/or time, Sex, Conditionality) to give predetermined logical outputs (after Labbe *et al.*, 2012). None of these genetic modules have been implemented in *Tribolium* previously. However, success with other species provides a plan of attack.

5.3.1 TRANSGENESIS IN *TRIBOLIUM*

Prior to these genetic modules we need a mechanism for detecting success of *piggyBac*-mediated transformation. Fortunately, *Tribolium* has been transformed on several occasions using a variety of promoter-marker combinations.

5.3.1.1 OX4700 DESIGN – IS A-TUBULIN-1-DSRED2 A SUITABLE TRANSFORMATION

MARKER?

Endogenous promoters were recommended (Schinko *et al.*, 2010); accordingly the promoter of *Tribolium castaneum* $\alpha 1$ *tubulin* (or α -*tubulin-1*) was used (Siebert *et al.*, 2008; Pfeiffer *et al.*, 2010). The construct was restricted to as simple a design as possible to demonstrate proof of principle in this species and corroborating the findings of Siebert *et al.* (2008) that $\alpha 1$ *tubulin* promoter provides constitutive expression throughout most tissues. Only one independent transgenic line was created so replication of their findings is appropriate before testing a more complicated construct design.

The construct OX4700 (Figure 89) was co-injected with a helper plasmid. Due to the sub-optimal transcription of transposase under control of exogenous promoters available it was preferred to use mRNA, sidestepping the transcription machinery of the cell and allowing immediate translation of the transposase (Lie *et al.*, 2012).

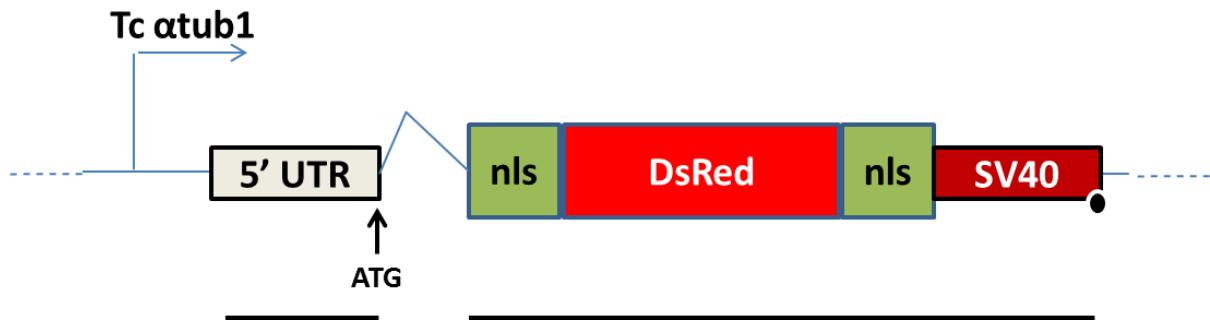


Figure 89. OX4700 schematic, designed for microinjection into *Tribolium castaneum* embryos to assess basic transformation marker functionality. The endogenous α -tubulin-1 promoter is designed to regulate expression of the marker DsRed2 and confer all-body DsRed2 fluorescence. To drive expression the promoter (359 bp), the 5' UTR (88 bp) and intron (154 bp) are required upstream of the reporter gene DsRed2 (675 bp). The nuclear localisation signal is found at either end of DsRed2 (48 & 36 bp). It has not been determined which end is necessary for functionality so nls are located either end of the protein to facilitate DsRed2 transport to the nucleus improving the conspicuousness of the marker module. The SV40 comprises the 3' UTR region at 228 bp in length providing transcript stability. The resultant mRNA is represented by the black line along the bottom, the wider rectangles will be translated into protein. The DsRed2 will be transported back to the nucleus due to the nls. The start codon is annotated with an arrow. The black blob represents a polyA signal on the SV40. Our design contrasts to Siebert *et al.*, (2008) in that we used DsRed2 instead of GFP as we did not have it on license.

Embryos were not screened for transient expression of the marker due to their sensitivity and the questionable usefulness of this as an indicator (Chapter 3). Larvae were screened when using a new injection mix to check for at least one instance of transient to verify designed marker protein expression.

5.3.1.1.1 GERM-LINE TRANSGENESIS (OX4700)

Each injection survivor was initially set up in single-pair crosses, with counterpart mates from the in-house GA-2 strain (wild-type), at the pupal stage. It was observed that some adult mortality was occurring before viable eggs were produced. As a precaution, insects were pooled in larger numbers as mortality can arise due to the beetles lacking a self-righting mechanism and subsequently starving. G_1 were screened for DsRed2 at the larval or pupal stage. Expression pattern of the *Tca-tubulin-1-DsRed2* reporter was expected to match that of Figure 5 in Siebert *et al.* (2008).

Table 26. The numbers of independent microinjections (n_1) of construct OX4700 into *Tribolium castaneum* embryos and related survival (G_0), transience (G_0t), number of offspring from crosses (G_1) and putative independent transgenic lines (X). Numbers in brackets give proportions of survival and transformation efficiency (2.d.p).

| Injections (n_1) | Survivors (G_0) | Transient (G_0t) | G_0 | Progeny screened (G_1) | Transgenic lines (X) |
|----------------------|---------------------|-------------------------|-------|-------------------------------|-------------------------|
| 3766 | 382 (0.10) | NA | | 13912 | 7 (0.02) |

The transformation efficiency of 0.018 transformants per G_0 was relatively low compared to other described *Tribolium* transformation (Lorenzen *et al.*, 2002; Lorenzen *et al.*, 2003). The apparently low efficiency may have been exacerbated by the relative difficulty of spotting a DsRed2 transgenic compared to the literature where a green fluorescent protein reporter was used.

5.3.1.1.2 PUTATIVE TRANSGENESIS

Putative transgenic G_1 were observed, they were aged and checked at a later larval stage to confirm apparent protein fluorescence was not associated with auto-fluorescence at a particular life stage (Figure 90). The phenotype was similar to that shown in Siebert *et al.* (2008). To confirm transgenic status the insects were individually crossed with wild-type mates. The phenotypes of the offspring were scored to assess for multiple insertions (Table 27).

Table 27. Putative transgenic G₁ insects derived from unique G₀ crosses were separated and labelled as different putative lines. One transgenic male was crossed with five wild-type females (or one female with three wild-type males if no females were produced by a particular G₀ cross). Data for lines A and B were lost. Numbers in brackets give proportions 2 d.p.

| Line | Transgenic | Non-transgenic | X ² (2.d.p) | P(2.d.p) |
|------|------------|----------------|------------------------|----------|
| C | 15 (0.43) | 20 (0.57) | 0.71 | 0.40 |



Figure 90. The transformation marker of OX4700 is functional in larvae (after second or third instar suggested). Putative transgenic OX4700 (right of picture) compared to two wild-type of similar larval age in white light (top) and screened for DsRed2 (bottom).

The phenotypes varied between putative transformants with some larvae demonstrating increased fluorescence above the legs of the second and third thoracic segments. This corroborates the first-described insertion of α -tubulin-1-GFP where expression was seen in the

wing and elytra discs (Lorenzen *et al.*, 2003). However, the α -tubulin-1-DsRed2 is less conspicuous compared to published descriptions of α -tubulin-1-GFP expression in *Tribolium* (Siebert *et al.*, 2008). Screening at pupae is not recommended as auto-fluorescence of exoskeleton and diet renders false-positives more likely. The marker appeared inconsistent prompting uncertainty in identifying transgenics (Figure 91).



Figure 91. The transformation marker of OX4700 is functional but inconsistent. Fluorescent 4700C pupae (right pupa) compared to a wild-type (left pupa) under white LED light (right column) and red filter (left column). Notice the flour trapped under the elytra that fluoresce for the wild-type pupa. The transgenic insects were chosen at random from a fifth generation cross. The wild-type was chosen at random from a pool of 10 wild-type insects that resembled the transgenic in white light. When screening it is useful to have reference wild-type and transgenics at hand for comparison as the marker is not particularly potent.

As no evidence of foci associated with the nuclear localisation signal patterning (Figure 91) was detected, molecular confirmation was desired. Insertion of the construct was confirmed by PCR (Figure 92).

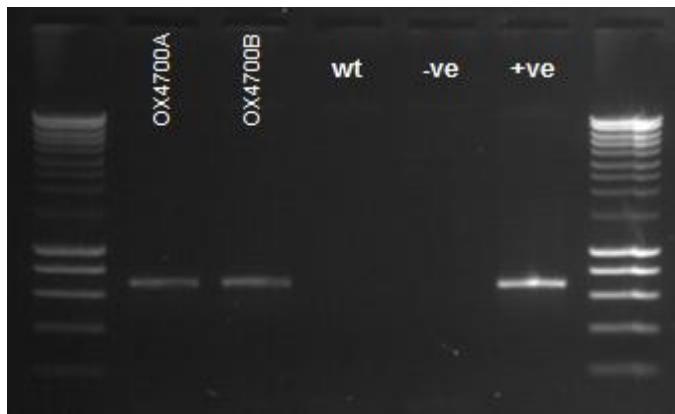


Figure 92. Confirmation of putative transgenic lines as testing positive for insertion events of the construct OX4700. Confirmed in *Tribolium castaneum* OX4700 lines by amplification of construct specific DNA sequence using primers (TD1406 and TD3026) across DsRed2 and the endogenous α -tubulin-1 promoter and 5' UTR. The expected band size was 659 bp. The negative (-ve) control is just primers, the positive (+ve) control - the miniprep of the construct diluted 1:100, was added in place of gDNA. The wild-type (wt) is negative for the target amplicon. Surrounding lanes provide DNA size standards. Fragment sizes are as follows: from bottom, 200 bp increasing in 200-bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

Bands of the expected size were detected for the putative transgenic insects suggesting successful insertion confirmed by sequencing and comparison with the designed construct.

Only the 5' region of the DsRed2 is shown for clarity. Interestingly the start codon is found upstream of an intron and following the 5' UTR of the *Tc* α -tubulin-1 promoter. The OX4700 marker matched the expected.

5.3.1.3 MARKER RELIABILITY

When screening the insects difficulty arose due to weak fluorescence at different life stages. This raised the question of how likely was it that transgenic insects could be missed due to the poor marker and could explain the relatively low transformation efficiency observed. The fluorescent status of an insect was compared to the genotype assessed using PCR (Figure 93).

A stock of OX4700C was used to supply the adult beetles (assumed to be heterozygotes and some homozygotes). The previous generation consisted of transgenic males and females with wild-type females added. Twenty-four insects were chosen at random by using carbon dioxide to immobilise and then generating a random number to chose. The insects were screened for

DsRed2 under a fluorescence microscope by examination of the ventral and dorsal side and putatively labelled positive or negative for DsRed2. The DNA of the insects was extracted and a PCR run with primers for an endogenous gene (molecular sex identifier; Methods 8.2.26) and the OX4700 transgene.

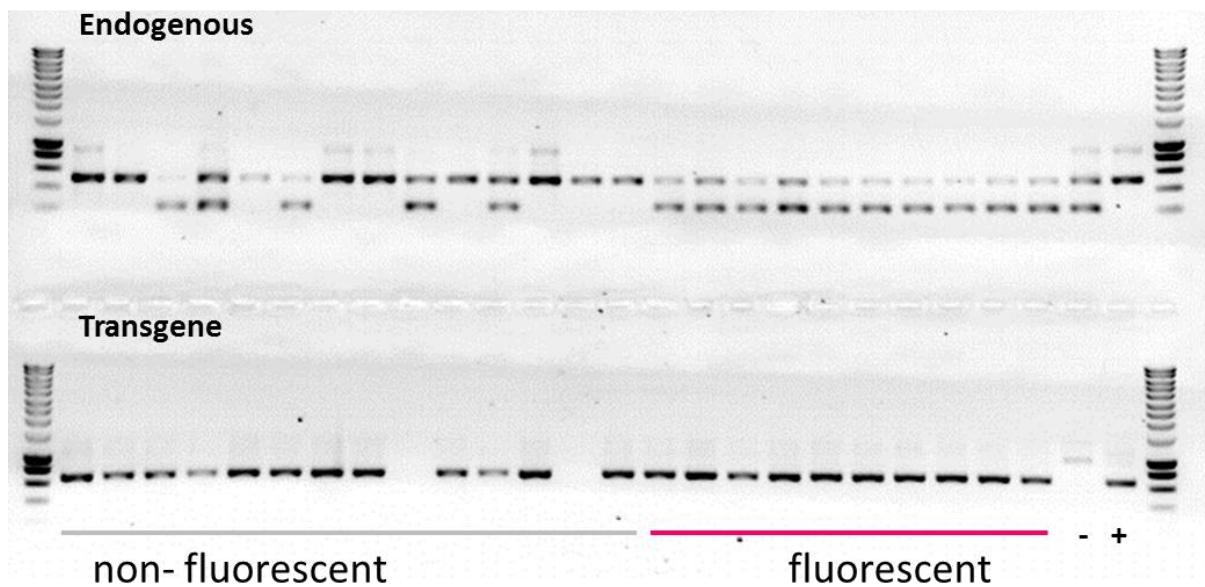


Figure 93. The transformation marker of OX4700 is poor with unreliable fluorescence microscopy assessment. A PCR was run using the extracted DNA of adult insects sorted by fluorescence. The endogenous control assessed the quality of the DNA and also provided a genetic means to assess the sex of the insects. Primers TD3372, 3373 and 3374 were used in a multiplex PCR with both sexes having a band at 500 bp and a male-specific band at 230 bp. Genomic possession of the transgene was assessed using primers across DsRed2, TD1406, and the endogenous α tubulin promoter and 5' UTR, TD3026. The expected band size was 659 bp. A negative control of a wild-type adult was used. The positive control was gDNA from OX4700Z. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

Interestingly all the insects that tested positive for fluorescence were male and 9 of 14 that tested negative were female. This suggests that DsRed2 may be more conspicuous in the males (as adults).

Only two of the non-fluorescent insects did not carry the transgene. All 10 of the fluorescent insects were positive for the transgene. In order to provide a quantitative measure of the validity of a test, in this case, the screening of insect for DsRed2, a two-by-two contingency table was constructed (Table 28).

Table 28. A contingency table summarising the marker reliability experiment where the validity of DsRed2 screening (either fluorescent or non-fluorescent) for transgenics (either transgene containing or non-transgene containing) was evaluated against a gold standard of DNA extraction and transgene status confirmation using PCR. The letters are provided to assist with term definition; where a) true positives, b) false positives, c) false negatives and d) true negatives. By chance most of the insects sampled were transgenic. The test correctly identified only 10 of the 22 transgenic individuals giving a sensitivity of 45% (a / a+c). The test was better on specificity (b / b+d) with no false positives although the low number of true negatives to test made it difficult to assess the reliability of this estimate. The positive predictive value (a / a+b) was 100%. The negative predictive value was 14% (d / c+d). The overall accuracy was calculated as 50% (a + d / N). However, this value does not take into account chance agreement. A test that is correct only half the time is no better than a coin toss. Cohen's Kappa was calculated to correct for chance agreement (0.43). Where $\kappa = [\text{observed agreement} - \text{chance agreement}] / [1 - \text{chance agreement}] = [0.50 - 0.43] / [1 - 0.50] = 0.12$.

| Test result | True status | | Total |
|---------------|-------------|----------|-----------|
| | transgene | non | |
| fluorescent | 10 (a) | 0 (b) | 10 |
| non | 12 (c) | 2 (d) | 14 |
| Totals | 22 | 2 | 24 |

With a Cohen's Kappa value of 0.12, even after correcting for chance agreement, we have a poor level of agreement almost equivalent to a coin toss (near 0); suggesting the marker is not fit for purpose. A caveat, this value can be misleading if the incidence rate is extreme (in this case there were only two true negatives of 24).

The results suggest that the marker may be more easily detectable in adult males. Furthermore, the test (screening adults under a fluorescent microscope) is poor at identifying transgenics, with half going undetected. This is a concern when screening for transformants following injections where the probability of getting transgenic G₁ is already low. It is likely that this could result in transformants passing through screening undetected and inappropriately discarded with half of them being missed. This suggests a new marker was required.

5.3.1.4 COMPARISON OF OX4700 AND OTHER PUBLISHED MARKERS

Siebert *et al.*, (2008) described the expression pattern produced by the α -tubulin-1 – EGFP to be:

“EGFP expression in the aT-EGFP line is ubiquitous, but in adults is particularly strong in the midgut epithelium, especially in the structures known as “crypts” or “papillae” that appear as eversions into the hemolymph. There is also strong expression in the hindgut epithelium, testes and testis accessory glands.”

Sequencing the insertion site showed the transgene of this single line to have been inserted the promoter region of an adipokinetic hormone receptor gene. In order to expand on the characterisation of this transformation marker we set out to determine tissue- and cell-specific marker expression patterns at higher resolution. This was particularly relevant to future development of a paternal-effect system, as any testes expression may interfere with our ability to detect sperm-marking. Furthermore, the *Tc* α -tubulin-1 gene is homologous to the widely used and well characterised *Dm* α Tub84B gene; elucidating expression patterns will reveal similarity between the dipteran and coleopteran model organisms.

5.3.1.4.1 ADULTS

5.3.1.4.1 MALE

The fluorescence in the adult male appeared to be towards the lower abdomen. The fluorescence at the adult stage was more conspicuous in the males compared to the females. The testes were visible through the cuticle of the male (Figure 94A).

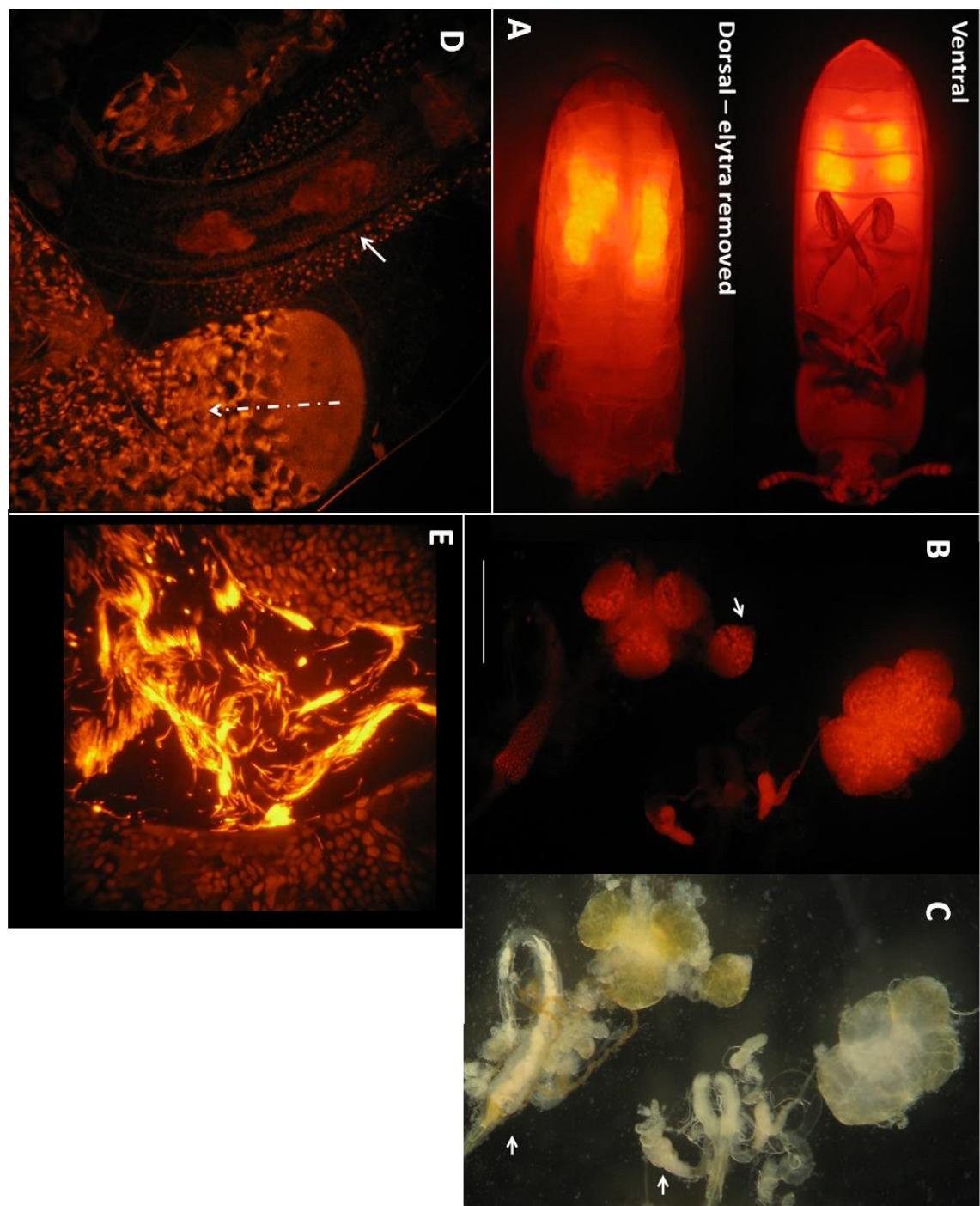


Figure 94. Investigating *a-tubulin-1*-DsRed2 expression in the OX4700C adult male reproductive system. (A) Testes appear to be the main source of DsRed2 fluorescence in *a-tubulin-1*-DsRed2 adult males. OX4700C male viewed ventrally (top) and dorsally (bottom). The elytra, wings, head and thorax were removed for the dorsal image. Image taken under DsRed2 excitation and filter conditions. (B & C) The reproductive organs that fluoresced when screened for DsRed2 included the testes (a lobe is shown on the left arrow), the putative accessory glands (right arrow) and the cells surrounding the ejaculatory duct. The white bar represents approximately 5 mm. (D & E) Images taken at x400 (left) and x600 (right) magnification showing DsRed2 fluorescence of male reproductive system anatomy. A testis lobe has been juxtaposed alongside the ejaculatory duct demonstrating nuclei fluorescence in both germ line and somatic cells (bold arrow). The dashed line shows an area of spermatogenesis whereby cells develop through a gradient from germline stem cells to spermatids. Sperm bundles fluoresce strongly compared to surrounding putative somatic cells in the testes.

Dissecting out the male reproductive system revealed the testes as well as other organs in the male reproductive system as responsible for the high DsRed2 expression (Figure 94B & C). This may explain why fluorescence was more easily detected in adult males for the marker reliability experiment.

Increasing the magnification and focusing on the testes revealed DsRed2 expression in sperm at different stages of spermatogenesis (Figure 94D & E). The putative somatic cells are darker in comparison.

This promoter-marker is difficult to detect in adult females. It will not be useful for development of a sperm marking system as the α -tubulin-1 driven transgene is expressed in sperm and bleeds through to green confounding sperm marking. This confirms and expands on Siebert *et al.* (2008).

5.3.1.4.1.2 FEMALE

Adult females showed auto-fluorescence in the gut and some fluorescence in the reproductive system. This is discussed in more detail later for construct OX5118 (5.3.5 OX5118, green transformants and red females).

5.3.1.4.2 PUPAE

5.3.1.4.2.1 MALE

In the pupae the male testes showed a high expression level of DsRed2 compared to other tissues (Figure 95). The testes under the fluorescent microscope proved as reliable a sexing tool as looking for the male-specific phenotype under white light (30/30 male pupae were correctly identified from a random sample of 100 hundred insects from the OX4700C stock population).

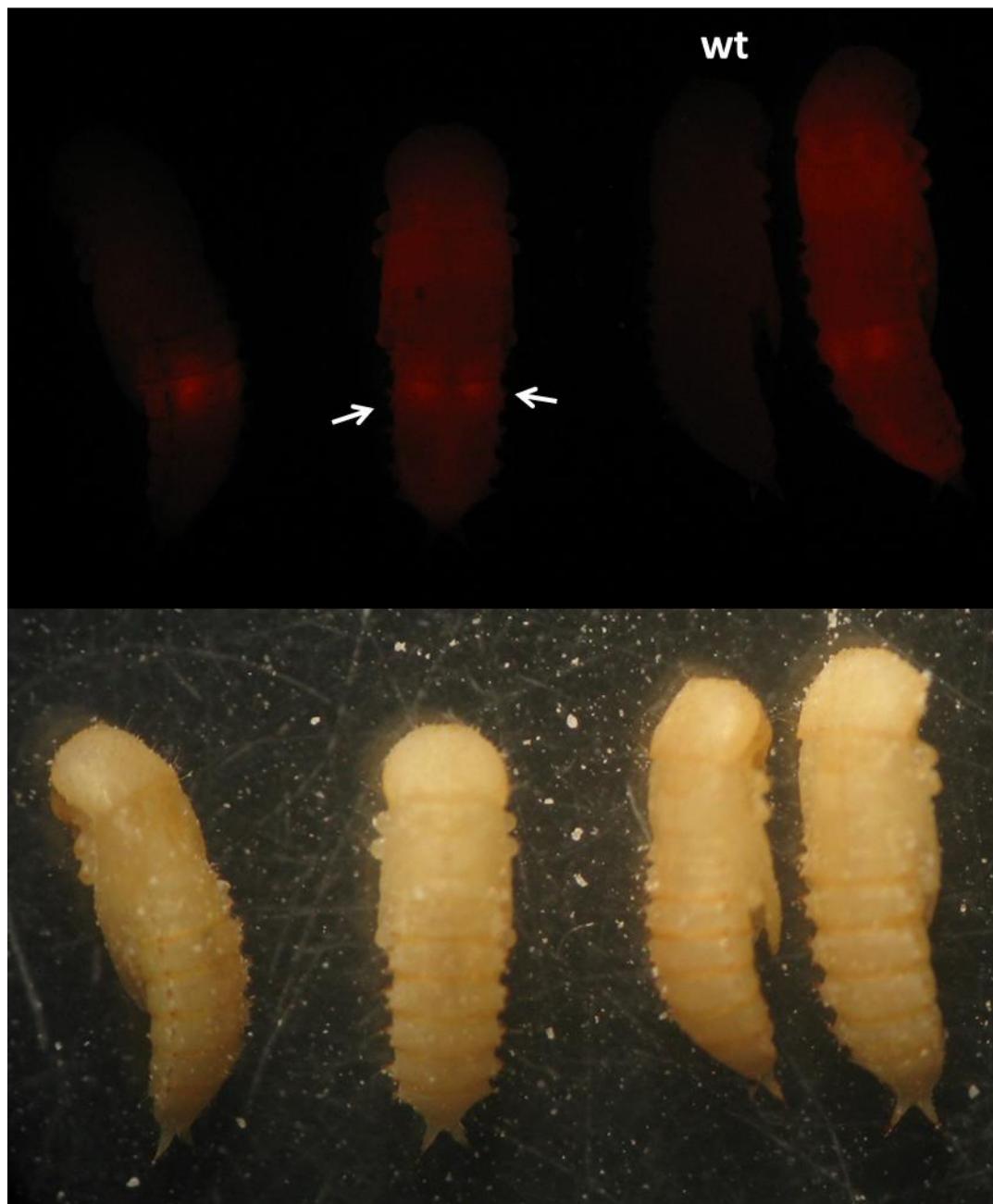


Figure 95. The transformation marker allows sex discrimination due to discrepancy in DsRed2 fluorescence. Four OX4700C male pupae are examined under red filter inspected for DsRed2 expression. Fluorescence is located mainly in the two testes (arrows). As the males mature the fluorescence becomes less obvious through the rest of the insect but remains visible in the testes. In comparison, the wild-type counterpart (wt) does not fluoresce. Mature transgenic female pupae lack this distinction so may be harder to distinguish from wild-type.

Each testis comprises six follicles. Within each follicle was a microcosm of spermatogenesis; from germ cells to spermatids. The α -tubulin-1 promoter drove expression of DsRed2 with a similar pattern (Figure 96) to that observed in the diamondback moth follicles when using *in situ* hybridisation of β 2-tubulin. Expression appears to start (DsRed2 takes hours to days to

mature) after cyst formation. Sperm differentiation follows a gradient towards the centre of the testis while remaining packaged in the cyst. The developing sperm are bundled as they continue to mature into spermatids. When screened for green no signal was detected.

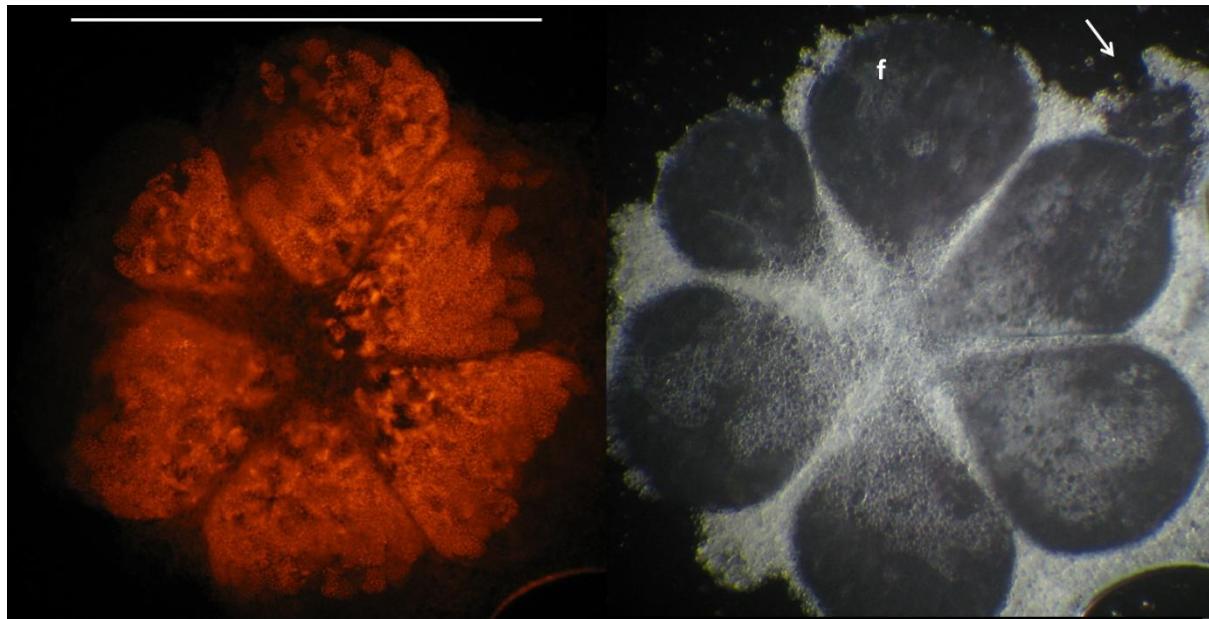


Figure 96. Six follicles (f) of an OX4700 *Tribolium* pupa testis are shown. Under a glass cover slip the testis is 900 μm in diameter (white bar). The images show DsRed2 and white light respectively. Dissecting the testis can result in damage as indicated by the arrowed tear; interestingly the cysts can be seen to move out of the follicle due to the negative pressure. Typically spermatogenesis continues along a gradient from the exterior to the interior. Each follicle is separated by the peritoneal membrane. The photo was taken at 100x magnification.

A close up of a single follicle reveals the timing of expression as probably occurring after the mitotic stages of spermatogenesis. This is relevant for future construct design of a paternal effect system in *Tribolium*. Any expression overlaps between the α -tubulin-1 and the putative *Tc* β 2-tubulin promoter during spermatogenesis may cause problems when identifying candidate promoters to drive nuclease-chaperone expression.

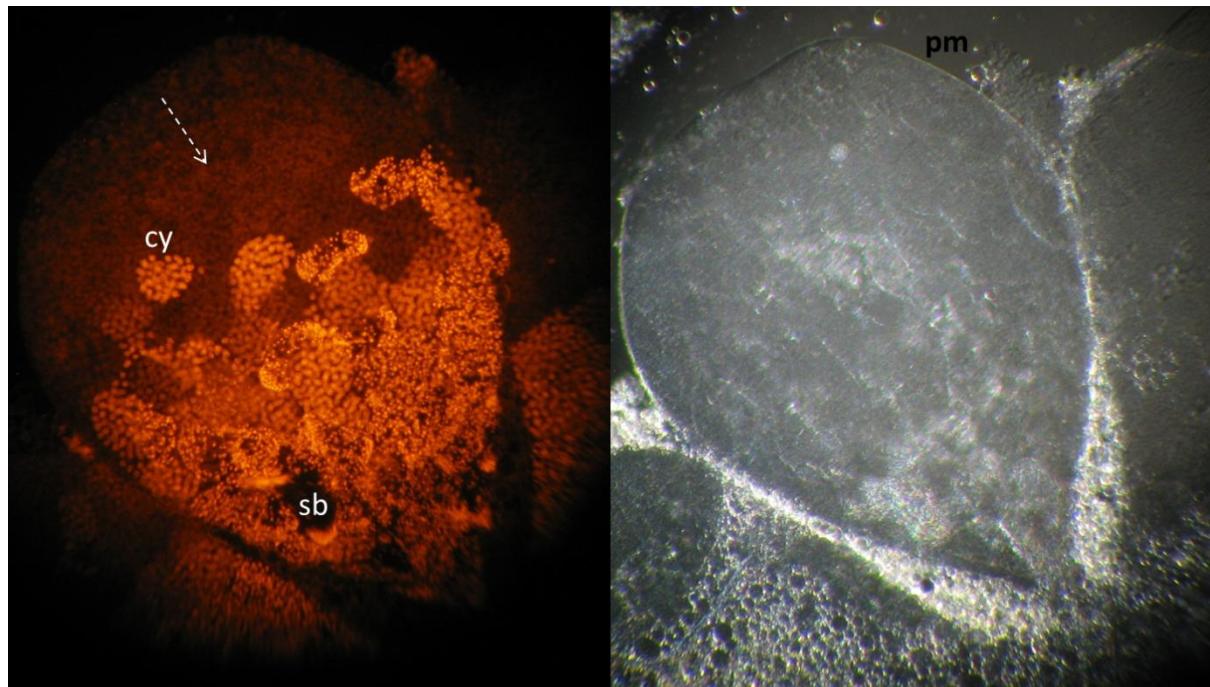


Figure 97. High magnification image of OX4700C follicle showing spermatogenesis. Taken at 400x magnification reveals additional detail of a single follicle and the timing of α -tubulin-1-DsRed2 expression. The first third of the follicle appears to be clear of DsRed2 (dashed arrow). A cyst (cy is shown) clearly enclosing > 30 cells. Sperm bundles (sb) are visible later shown by longitudinal smears of concentrated DsRed2. Late spermatids are also visible albeit not clearly at this resolution. Each follicle is separated by the peritoneal membrane.

5.3.2 REARING *TRIBOLIUM* ON A TETRACYCLINE-DOSED DIET

RIDL technology is dependent on conditional transcription regulation of a gene of interest being regulated by tetracycline presence or absence. Transgenesis has been achieved in *Tribolium* but for the specific requirements of RIDL technology it is important to determine whether the beetle can ingest tetracycline and survive thereafter. *Tribolium* is unusual relative to many other insects reared in the laboratory on artificial diet, in that it lives and feeds in a dry, particulate diet. This presents a potential problem in terms of mixing the antibiotic evenly into the food and ensuring that the beetle ingests a sufficient quantity as an antidote to regulate transgene-induced lethality.

A bioassay was conducted to gauge the effect of tetracycline (Chlortetracycline Hydrochloride, ELS, 26430 -100G) on *Tribolium* mortality or survival (Methods 8.1.5). An excess of diet was mixed with a ratio of 95% wholemeal wheat flour to 5% brewer's yeast. The factor of

tetracycline concentration contained four levels. These concentrations were chosen as they provided a negative control (i), a comparison with a concentration typically used to rear RIDL diamondback moth (ii) and two extreme concentrations different by a factor of ten (iii and iv). Thus they provide a scale against which a comparison can be drawn to other insects reared on tetracycline (Harvey-Samuels, 2014).

- i. Control (0 ppm or 0%)
 - ii. 0.2 µg/mg (200 ppm or 0.02%)
 - iii. 10 µg/mg (10,000 ppm or 1%)
 - iv. 100 µg/mg (100,000 ppm or 10%)

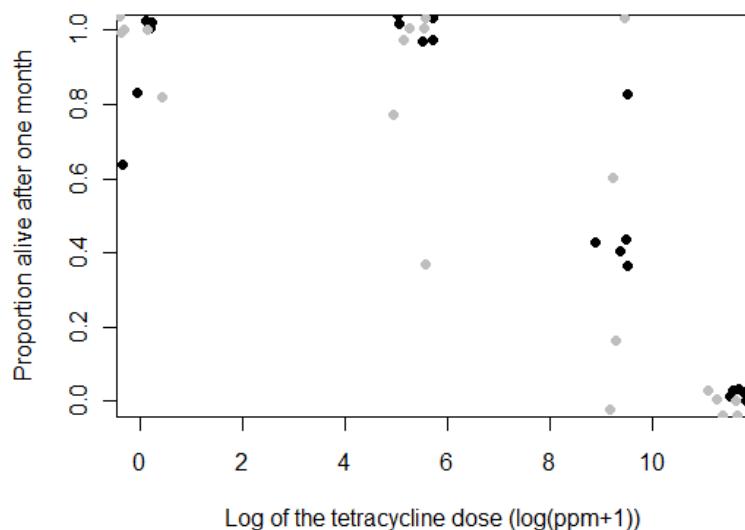


Figure 98. Tetracycline does not affect survival at typical doses used to suppress RIDL phenotypes in other insects. The survivorship after 1 month of *Tribolium castaneum* fourth-instar larvae (grey) and 3-week-old adults (black) reared on diet (95 wheat flour : 5 brewer's yeast) with tetracycline added. For comparison, the typical Lepidopteran diet would be at 4.6 (log (100 ppm+1)). Data has been "jittered", horizontally and vertically, to preclude overlap of the data points ($n = 40$, at four different concentrations of tetracycline, 0, 0.01, 0.1 and 0.5).

A generalised linear model with quasibinomial error distribution was used to describe the proportion alive after one month. The model had an overdispersion factor of 1.2, model

simplification went ahead from the maximal model using the F test to compare models and check for any reduction in explanatory power. The interaction term and factor of life stage were found to be unimportant in describing the data, with no significant loss in explanatory power when removed from the model ($F = 0.0693, p = 0.79$; $F = 0.207, p = 0.652$). This indicates that survival was consistent between life stages.

The one-month-mortality by concentration was plotted and the LC50 was estimated to be at 16.8 $\mu\text{g} / \text{mg}$ or approximately 1.6% of the diet as tetracycline by mass (using the drc package in R; Ritz & Streibig, 2005) (Figure 99).

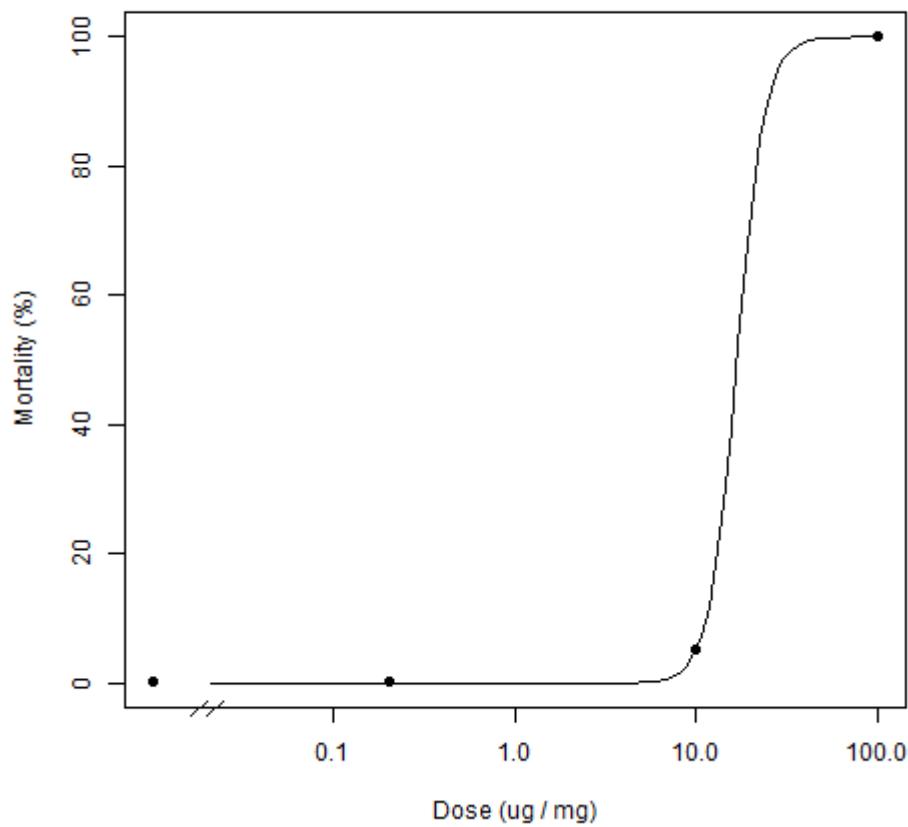


Figure 99. The dose-response curve of *Tribolium* larvae reared on a tetracycline-supplemented diet after one month. The resolution between doses 10 and 100 $\mu\text{g}/\text{mg}$ was lacking, hence a large standard error for the LC50 of 62 $\mu\text{g}/\text{mg}$.

Tetracycline concentration in the diet has an effect on survivorship but is only a concern for rearing when using very high concentrations. At typical operating concentrations it was shown

to be unlikely to affect *Tribolium* adult survival or metamorphosis of the insect at the sensitive fourth instar and pupal stages. However, in other insect diets, tetracycline is delivered dissolved in solution prior to solidifying, whereas in *Tribolium* diet it is mixed in powdered form and may be less homogenously distributed.

5.3.2.5 TETRACYCLINE DELIVERY TO MALE GONADS

Ultimately the paternal-effect system envisioned by this thesis and set out in the introduction will require tetracycline to be delivered to the testes to repress tTAV activity and the resultant paternal-effect sperm-DNA damage during mass-rearing of the insect. It should be determined that tetracycline can be administered to the insect through its diet and delivered to the appropriate tissues.

Wade and Stevens (1985) showed that tetracycline applied to the diet at a concentration of 0.125% (or 0.0625 g of tetracycline per 50 g of medium or 1250 ppm) “cured” the *Tribolium* of *Wolbachia*, a micro-organism found in some cells (notably the gametes). This indicates tetracycline can be delivered to the sperm regulating a repressible paternal-effect system.

A similar experimental setup and procedure to that described above was used except with first instar larvae (1-2 day old since hatching from embryo). Each replicate consisted of five randomly chosen first instar larvae placed into a small Petri dish with 8 g of 95:5 wholemeal flour and yeast. Ten replicates had tetracycline added ten did not. The number of extant individuals was recorded every few days for over a month. The proportions of extant insects were compared between treatments to determine a treatment effect if any.

5.3.2.5.3 THE RESULTS

Not all the insects survived to adulthood, as mortality is relatively high during development through the larval stage. The experimental units exposed to tetracycline-enriched diet were more variable than the non-tetracycline control. The median survival for the 10 experimental

units per treatment were both 100% (Figure 100). Pearson's chi-square test (Howell, 2010) on the data revealed no significant difference between treatments on the survival of first instar larvae ($X^2= 0.0073$, $df = 1$, $p\text{-value} = 0.0073$). The data was plausible given the null hypothesis of no difference between treatments. The more conservative method of simulating p values 10,000 times using the Monte Carlo method reinforced this message ($X^2= 0.0535$, $df = NA$, $p\text{-value} = 0.887$) (Hope, 1968).

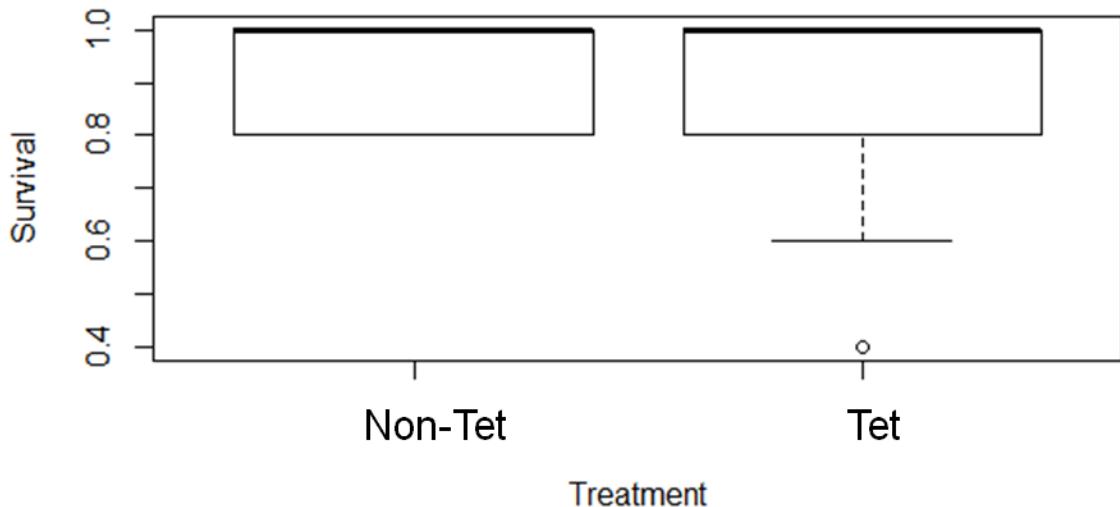


Figure 100. Effect of tetracycline on first-instar larvae survival. Box and whisker plot of the distribution of the derived variable, survival (number of insects developed into adulthood after 40 days). Each experimental unit consisted of five insects in a pot; accordingly survival has a finite distribution of possible values, ten replicates per treatment. Both median and lower quartile ranges are identical. The tetracycline diet exposed replicates showed more variation with an outlier (more than 1.5 times the inter-quartile range from the lower quartile range) at 0.4. Ten replicates in total mitigated the outlier effects in statistical analyses using conventional comparisons described in the main text.

Visual inspection of confidence interval overlap confirms conclusions derived from statistical tests (Figure 101). Given the number of replicates and the nearly bounded position of the mean the sensitivity of this experiment would only be able to detect differences approaching 0.20 or greater (simulation not shown).

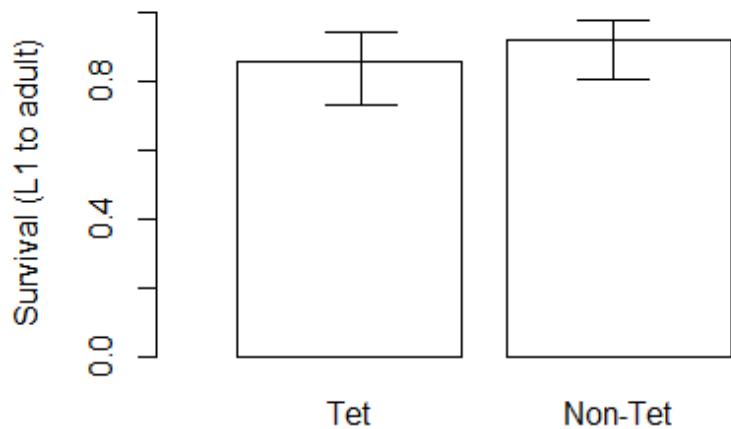


Figure 101. No effect of tetracycline on first instar larvae survival including variability statistic. The mean survival (the proportion of insects that developed into adults within 40 days) by diet of first instar larvae. Each replicate consisted of a count of the number of larvae reaching adulthood after 40 days relative to the five insects placed into each Petri dish at the start of the experiment. The 95% confidence intervals are shown for the plausible range of values that the mean could be given the data (Clopper & Pearson, 1934).

The data supports previous findings in this thesis relating to fourth instar larval with a potentially more sensitive life stage. Previous work by Wade & Stevens (1985) showed that tetracycline diet at this concentration (1250 ppm) had a phenotypic effect on Wolbachia infection, suggesting tetracycline ingestion and physiological interaction in the gonads. This is an important pre-cursor to the desired paternal-effect phenotype to be developed in the Coleoptera.

I have shown that *Tribolium* can survive on tetracycline enriched diet and that the tetracycline may be reaching the gonads. However, in order to test this hypothesis we require a genetic construct designed to demonstrate tetracycline-repressible lethality.

5.3.3 DESIGN OF A GENETIC SEXING CONSTRUCT (FEMALE-LETHAL) –

DOUBLESEX OR TRANSFORMER?

Research into sexing strains that exploit differential regulatory pathways and mechanisms in the sexes is abundant in the literature. *dsx* and *tra* are genes that both play a role in sex determination and both have homologues in many insect species.

5.3.3.1 DOUBLESEX

This work was completed with the assistance of Dr Tarig Dafa'alla.

5.3.3.1.1 PRIMER DESIGN FOR DOUBLESEX RT-PCR

The genome of *Tribolium* was sequenced in 2008 (Richards *et al.*). The genome was analysed for the presence of the *dsx* gene using Vector NTI and Basic Local Alignment Search Tool (BLAST).

A predicted nucleotide sequence was attained RT-PCR used to amplify the predicted mRNA of the *dsx* gene in both male and female *Tribolium* pupae (Figure 102). There were four bands of interest in total, each of which was cloned and sequenced.

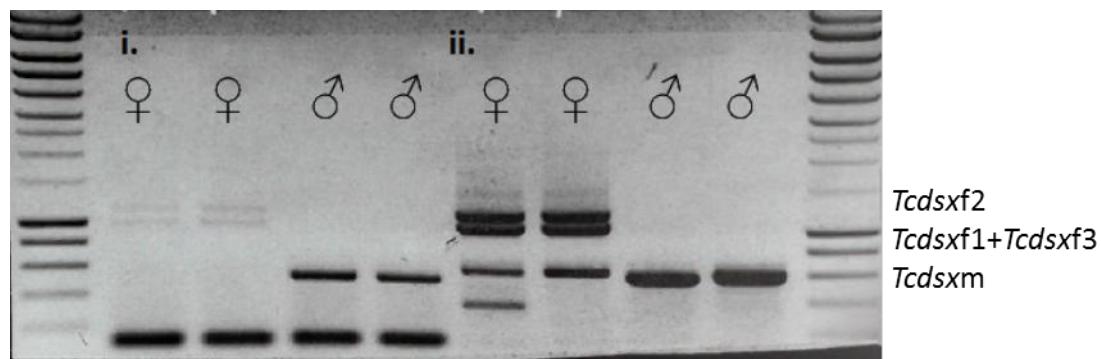


Figure 102. One-step RT-PCR with primers across *Tribolium castaneum doublesex* gene. Only two of the three female isoforms were readily discernible probably due to their similar sizes. Primer set i) TD3000 & TD3006 (545 bp), ii) TD3000 & TD3007 (570 bp). Primer dimers formed for reaction i but were ignored. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The sequence was interpreted using Vector NTI. Analysis showed that the single amplified fragment in the males resulted from *dsx* exons 2 and 4, both with a length of 141 bp. These exons were also shared by the *dsx-f* isoform (Figure 103). Exon 3 was female-specific and spliced out of the *dsx-m*. Exon 3 was shown to be two discrete exons (designated 3a and 3b respectively) with an intron of 164 bp between them.

This supports the findings of Shukla & Palli (2012) where it was reported that the pre-mRNA of *Tcdsx* is sex-specifically spliced to produce three female and one male-specific isoforms. All three female-specific *Tcdsx* isoforms were generated as a result of alternative splicing within the female-specific exon (exon3).

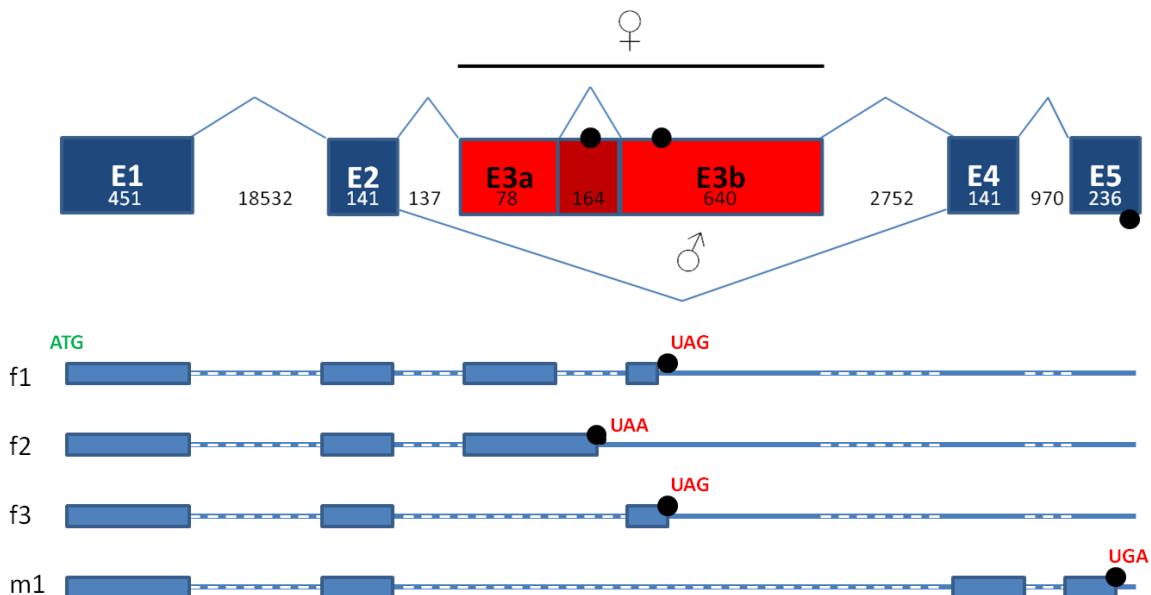


Figure 103. *doublesex* alternative splicing vindicated. Alternate splicing of *doublesex* in *Tribolium castaneum* females and males. Exon 3 is female-specific consisting of two discrete exons with a separating intron of 164 bp in length. The details of the different isoforms are shown where the blue boxes represent translated regions, the white dashes represent introns and the blue lines untranslated regions. Some sequence prior to exon 1 and after exon 5 is omitted. The proposed *Tcdsx* minigene contains sequence from exon 2 to exon 4 including the introns.

This data permitted the design of a construct with *Tcdsx* as the basis for a genetic-sexing system. Exons 1 and 5 were removed, as were the larger introns, to reduce the size of the insert. The *Tcdsx* minigene (Figure 103) was proposed to work by incorporating an effector gene into the minigene exon 3. This would provide female-specific expression as it would be spliced out of the

male pre-mRNA. I was ready to synthesise the construct when results from another group presented an alternative approach.

5.3.3.1.2 TRIBOLIUM CASTANEUM TRANSFORMER

The successful sequencing of *tra* in *Tribolium* gave us pause (Shukla & Palli, 2012). The structure of *tra* seemed more appropriate for genetic sexing compared to *dsx*. This was due to the male *dsx* stop codon being located in exon 5 of the male, due to size constraints it could not be included in our transgene, therefore any pre-mRNA not correctly spliced could result in transgene expression in the male, albeit at a low rate.

tra contained preferred alternative splicing isoforms between the sexes, with one female-specific isoform was identified; with knockdown by RNA interference resulting in male offspring (Shukla & Palli, 2012). The female isoform had a unique intron which could provide a means for female-specific expression (Figure 104). This was preferred to *dsx* as the stop codon containing exons were spliced out of *tra-f* yet maintained in the *tra-m* providing a simpler route to fsRIDL. Additionally, *tra* is upstream to *dsx* in the sex determination cascade.

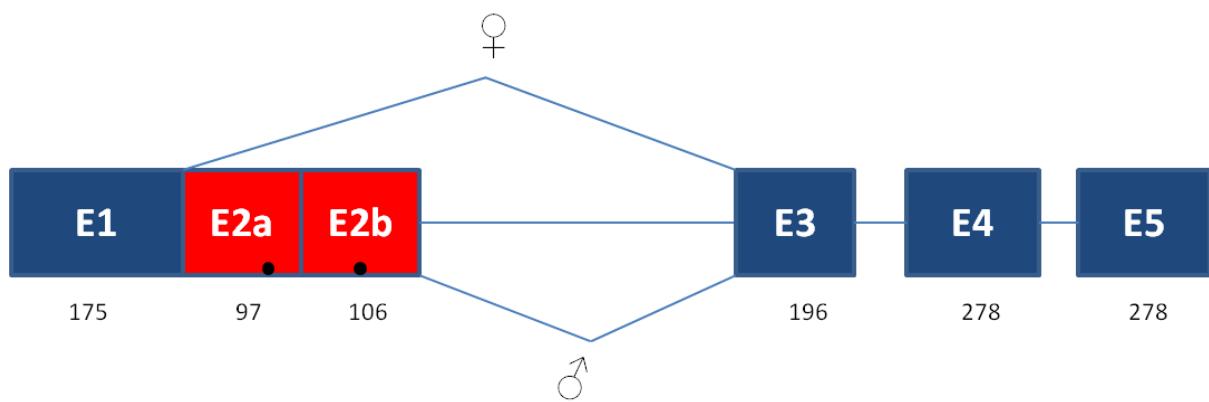


Figure 104. *Tctra* exon and intron configuration provided by Shukla and Palli (2012b) Common exons are shown in blue whereas male-specific exons are in red. Black circles represent stop codons. The sex-alternate introns are provided by the wavy blue lines with the common exon straight.

This work was verified using one-step RT-PCR on our wild-type (Figure 105). The amplicons were sequenced and compared to the published sequence. The sequences were sufficiently similar to proceed with development and design of a genetic sexing system exploiting the alternative splicing in *tra*.

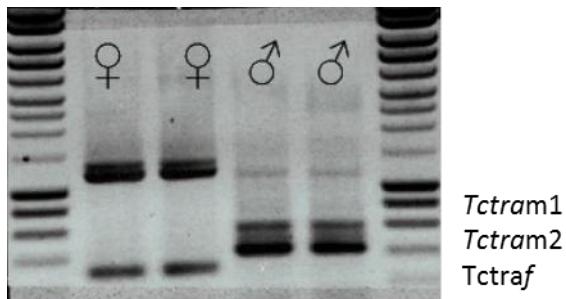


Figure 105. Confirmation of alternative splicing of *transformer* in *Tribolium* providing a path to fsRIDL. The amplicons of one-step RTPCR across *Tribolium castaneum* *transformer* with primers TD3184 and TD3187. These primers span from exon one to exon three, across the male exons and the shared intron. The single female isoform (*Tctraf*) is about 300 bp in size compared to the larger male isoform amplicons between 400 and 600 bp. According to Shukla & Palli (2012) the m1 and m2 isoforms should be 97 and 97+106 bp larger than f as females splice out the male exons. The band between 1 and 1.5 kb is likely gDNA contamination and / or non-specific amplification. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp. (Eurogentec, Smartladder).

5.3.4 OX4799 – TRANSGENE-BASED, FEMALE-SPECIFIC LETHALITY-BASED SYSTEM FOR GENETIC SEXING IN *TRIBOLIUM*

OX4799 was designed to confer lethality when transformants are reared off tetracycline, through female-specific transcription of tTAV. Following splicing of the transcript, males express non-functional protein. The females produce functional tTAV interacting with *tetO* and enhancing the *hsp70* minimal promoter and tTAV expression, creating a positive-feedback loop leading to death of the insect. The addition of tetracycline to the diet provides an antidote with tTAV preferentially binding the tetracycline over the *tetO* (Figure 106).

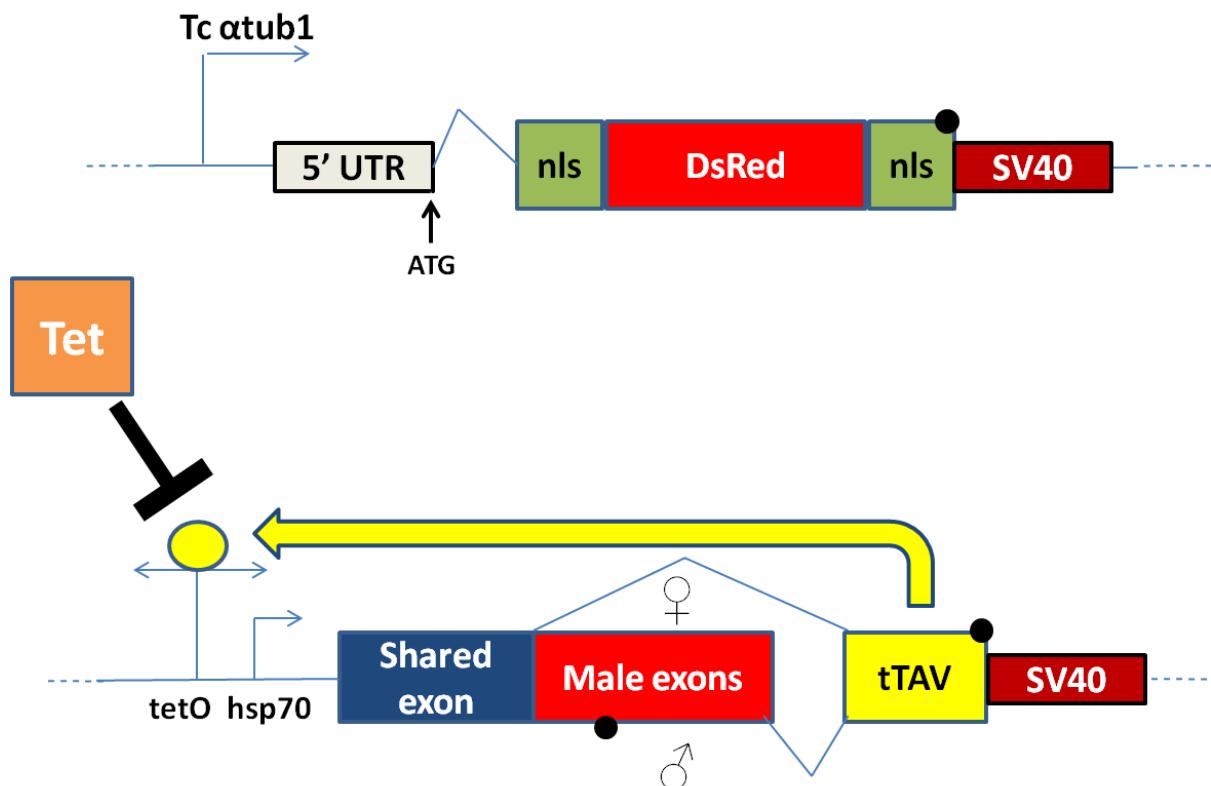


Figure 106. Schematic of proposed fsRIDL system in *Tribolium* using alternative splicing of *transformer*. An exploded view of the female-specific lethal construct OX4799 contains a conditional female-specific lethal gene (bottom half) and a fluorescent marker gene (top half) based within a *piggyBac*-based non-autonomous transposon vector (omitted for clarity). The non-autonomous *piggyBac* vector allows germ-line transformation of insects via microinjection with separately supplied *piggyBac* transposase. The fluorescent marker is DsRed2 under the control of an endogenous-derived promoter; *a-tubulin-1*, providing fluorescence. The conditional female-specific gene is designed to derive its lethal effect via accumulation of tTAV from a positive feedback loop, leading to hypothesised transcriptional suppression. Lethality is repressed by dietary tetracycline, which inhibits the binding of tTA to *tetO*. Female-specificity is provided by alternative splicing of the male and female transcripts leaving the female with functional tTAV and the male not. In the female the male exon with the stop codon (black filled circle) is spliced out of the transcript producing tTAV. The male transcript contains the male exon and the stop codon, preventing tTAV translation.

5.3.4.1 OX 4799 INJECTION DATA

Microinjection of the construct was conducted as standard (Table 29). Most of the mortality occurs during or shortly after microinjection of the embryo (90%) with further losses from larva to adulthood (15%) having a lower impact on the overall survival. Injection survivors were pooled and allowed to mate (pools varied from four to twenty). Sex ratios were female-biased to increase the number of potential matings. If inadequate numbers of G₀ females were available then pools were supplemented with wild-type females.

Table 29. OX4799 injection data. G₀ are injection survivors at adulthood used for crosses to generate G₁ offspring to be screened for DsRed2. G₁ that are positive for the transgene and from a unique cross are designated as independent insertions (X).

| Injected embryos | Larvae | G ₀ | G ₁ | X |
|---------------------|--------|----------------|----------------|-----------|
| 2007 | 370 | 298 (0.15) | 5041 | 1 (0.003) |

Of the first month-long batch of injections, three putative transformants were found at the larval stage. Two of the insects died at this stage — designated 4799A1 and 4799A2 — with the sex undetermined and the OX4799A3 died as a pupa. DNA was extracted from the pupa and its sex determined (Methods 8.2.26). The PCR product was run on a gel and produced bands of the expected sizes for a male.

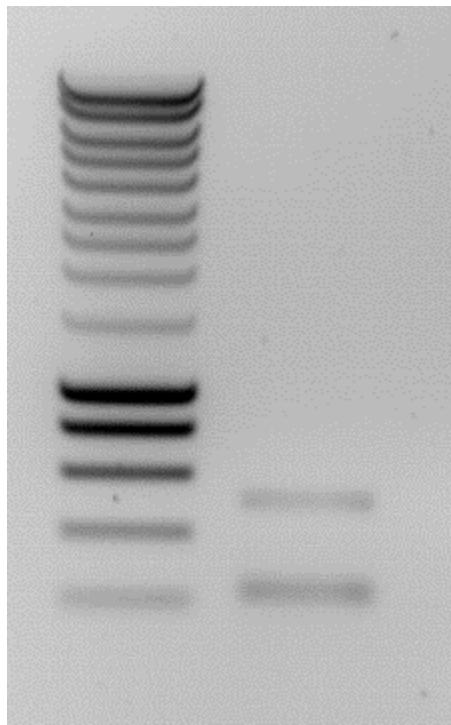


Figure 107. The putative transgenic OX4799A3 pupa was a male: determined by presence of two bands of approximately 500 and 230 bp respectively. The multiplex sexing PCR used primers TD3372, 3373 and 3374. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

To test for a successful insertion a PCR was conducted on the gDNA (Figure 108). The OX4799A amplicons were cloned, screened and sequenced; confirming insertion.

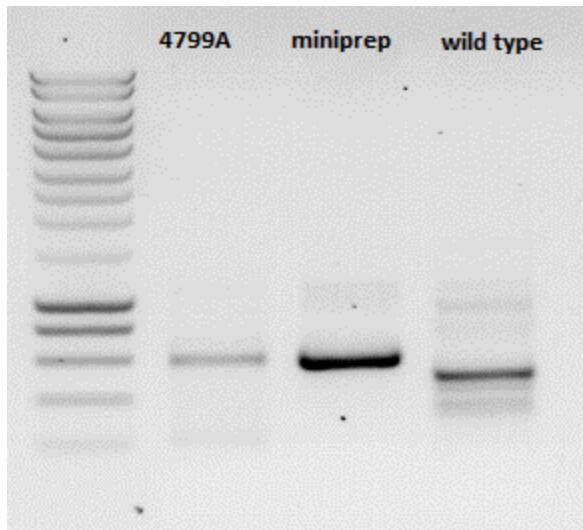


Figure 108. OX4799A and the miniprep test positive for the OX4799 transgene with PCR. The amplicon region extended from the *a1tubulin* promoter to the Tc *transformer* intron in the construct. The expected size of the amplicon specific to the sequence of the OX4799 construct was 676 bp. The band for the 4799A and diluted miniprep are a similar size but different concentration. The amplicon produced in the wild-type reaction appears to be smaller in size. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder). Bands of the expected size were detected (expected 676 bp) for the putative transgenic insects suggesting successful integration of the construct into the germline of *Tribolium* (albeit for only one generation).

5.3.4.3 A MORE CONSERVATIVE CONSTRUCT DESIGN

The total number of G₀ crossed suggests that more than one independent transgenic line should have been achieved assuming a similar transformation rate to that achieved for OX4700, even if all of one sex were to die due to construct-induced toxicity (the alternative splicing of functional tTAV in females). Given previous experiments with tetracycline and *Wolbachia* (Wade & Stevens, 1985) one would expect the concentration of tetracycline used in the *Tribolium* diet to be sufficient to rescue transgenics from the positive feedback of the *tetO* and tTAV RIDL module of OX4799 (Harvey-Samuel, 2014).

5.3.4.3.1 CONSTRUCT SIMPLIFICATION

In a re-designed construct, the *tetO* and tTAV components were removed from the genetic sexing module. Instead a fluorescent marker protein was added so that the females would have an altered phenotype relative to the males. DsRed2 was chosen for this role, with the previous transformation marker, DsRed2, replaced with a fluorescent green protein. ZsGreen is a green fluorescent protein option and has a proven track record in diverse insect species (Martins *et al.*, 2012). The nls was maintained on either side of the DsRed2 coding sequence for convenience of construct building.

The α -tubulin-1 promoter was designed to drive *Tctra*-regulated DsRed2 transcription. To mitigate the effect of recombination, its orientation was engineered to match that of the α -tubulin-1 sequence in the transgenesis module. The construct was designated OX5118.

5.3.5 OX5118, GREEN TRANSFORMANTS AND RED FEMALES EXPECTED

The modifications to the construct described above removed some of the complexity. Transformants were designed to fluoresce green on screening. The sex-alternate splicing of the *Tctra* was designed to produce DsRed2 in females only (Figure 109).

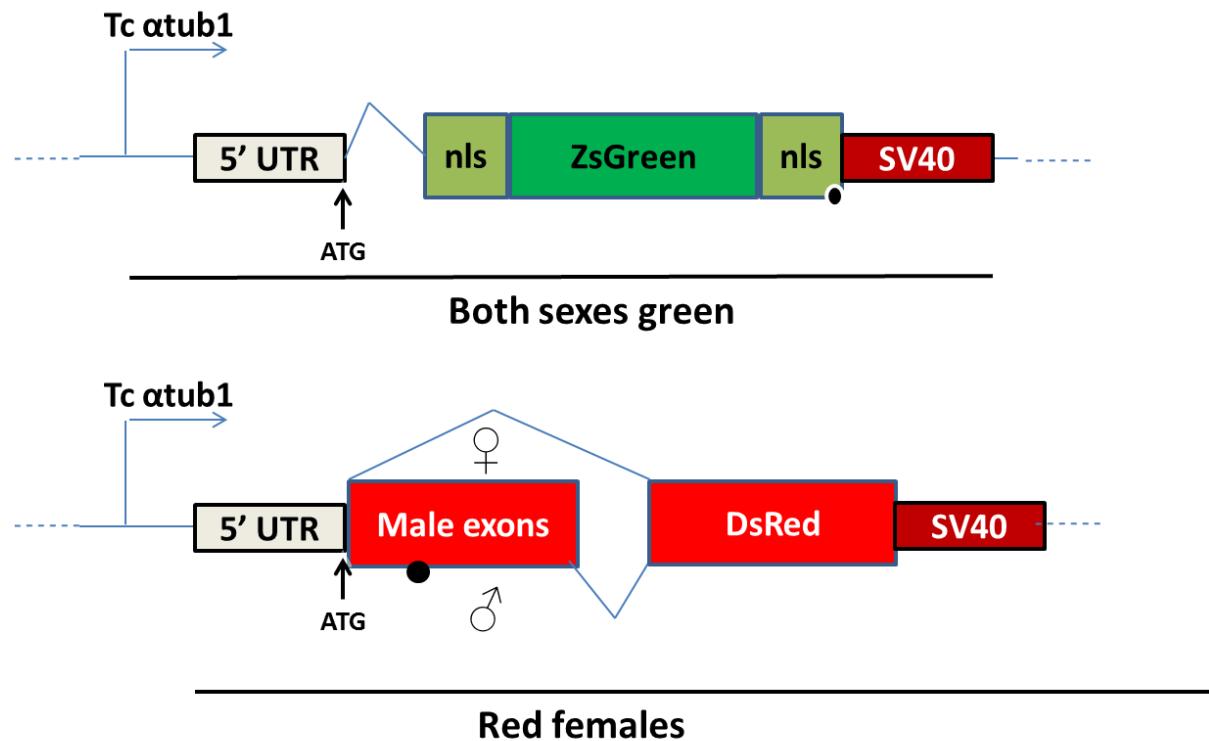


Figure 109. OX5518 schematic with female-specific expression of DsRed2 in *Tribolium* regulated by alternate splicing of *transformer*. For the transformation marker, DsRed2 was replaced with ZsGreen as it was predicted to be a superior reporter in *Tribolium* (top). The female-specific component was dependent on alternative splicing between the sexes (bottom). The females would splice out the male exons and stop codon therein producing functional DsRed2 protein.

The *tra* shared intron upstream from the shared exon was removed. The nls was removed from the DsRed2 to maintain compatibility between the adjacent sequence to the female intron and the putative splice recognition site.

5.3.5.1 INJECTION DATA

The construct OX5118 and helper mRNA, as part of an injection mix, were microinjected into embryos as standard (8.3.4.2) (Table 30).

Table 30. Injection data for OX5118 injected into *Tribolium* embryos. G_0 are the number of fertile adults derived from the injected embryos. G_1 were not counted to save time. The number of independent transgenic lines, X (conservative – only those from unique crosses were considered independent lines).

| Injected embryos | Larvae | G_0 | G_1 | X |
|-----------------------------|---------------|-------------------------|-------------------------|----------|
| 1327 | 200 | 154 (0.11) | - | 2 (0.01) |

5.3.5.2 TRANSGENESIS

A transgenic G_1 was identified at the larval stage and the marker was much more conspicuous compared to that of OX4700, which used DsRed2. Two putative-unique lines were identified. Both sexes were negative for DsRed2 fluorescence. Each transgenic was crossed with wild-type of the opposite sex. The offspring were reared and scored for transgenic: wild-type ratios using fluorescence microscopy (Table 31). One-to-one ratios were detected indicating a single insertion of OX5118 in each line. Screening was simplest with larvae or pupae (Figure 110). Adults had to be viewed ventrally for confirmation of transgenic marker fluorescence.

Again it was noted how the darkening of the elytra made the older pupae less conspicuously fluorescent (Figure 111). However, they were still readily identifiable under the fluorescent microscope and a vast improvement on DsRed2 as a reporter.

Table 31. Putative transgenic G_1 insects derived from unique G_0 crosses were separated and labelled as different putative lines. One transgenic male was crossed with five wild-type females (or one female with three wild-type males if no females were produced by a particular G_0 cross). The frequency of transgenic to non-transgenic insects after one month three weeks were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios (or a proportion significantly greater than 50% transgenic (indicated by EGFP fluorescent protein marker) to non-transgenic phenotype). Evidence for multiple insertions of the transgene in any line was weak. Proportions of the total shown in brackets.

| Line | Transgenic | Non-transgenic | χ^2 (2.d.p) | P (2.d.p) |
|-------------|-------------------|-----------------------|------------------------------------|------------------|
| B1 | 79 (0.53) | 71 (0.47) | 0.43 | 0.51 |
| A4 | 91 (0.50) | 90 (0.50) | 0.01 | 0.94 |

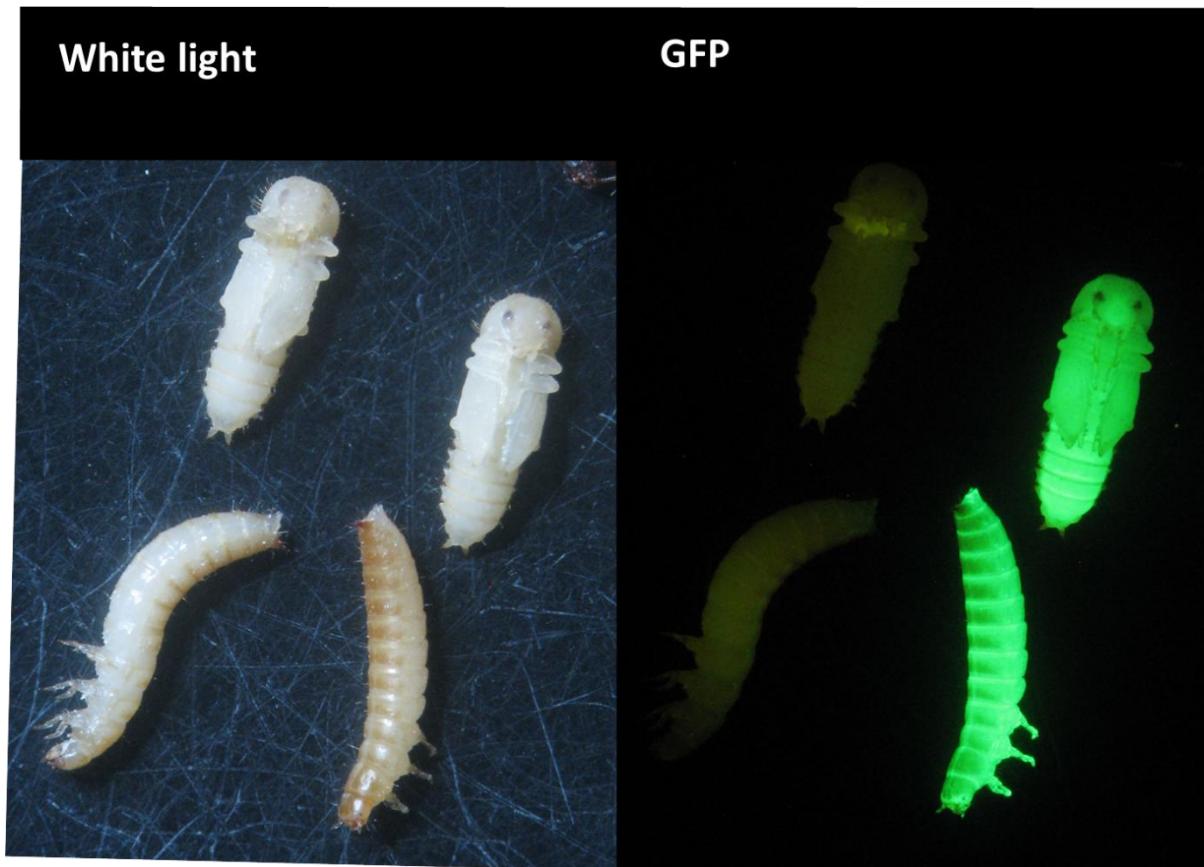


Figure 110. *a1-tubulin-ZsGreen* shows strong expression in pupae and larvae. OX5118 positive (right column of insects) *Tribolium castaneum* pupae and larvae compared to wild-type (left column of insects) in white-light (left image) and under fluorescence microscope with GFP filter (right image). The transgenic insects fluoresce conspicuously compared to the wild-type under GFP imaging.

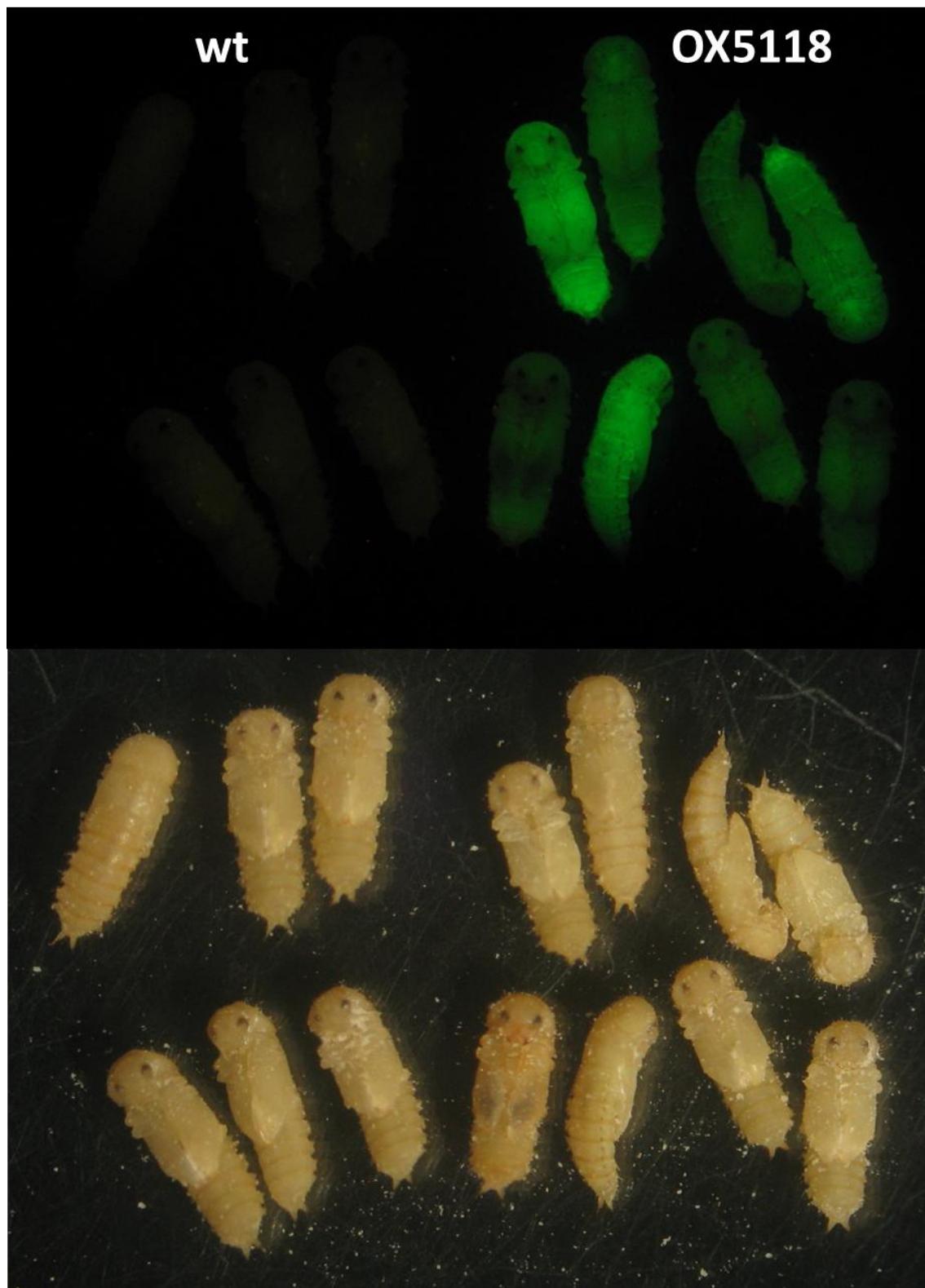


Figure 111. A description of the phenotypic range of Zsgreen conspicuousness in OX5118A4-positive pupae. The difference between the wild-type (wt) and the transgenic is more obvious when screened by eye than captured by the camera. This image also shows the phenotypic range of ZsGreen expression. The luminosity probably decreases as the pupae mature and their elytra darken due to melanisation. The top half is screened using a GFP filter and the bottom is under white light.

Adults could be screened after cooling or exposure to CO₂ to retard locomotion. The beetles were flipped and examined dorsally for fluorescence at the joints between chitin plating. Adults with a pale exoskeleton were more easily screened (Figure 112).



Figure 112. *a1-tubulin-ZsGreen* expression in adults is inconspicuous in areas with darkened exoskeleton. Adults are more difficult to screen requiring flipping to view ventrally. OX5118 positive (top two adults) *Tribolium castaneum* adults compared to wild-type (bottom insect) in white-light (right image) and under fluorescence microscope with GFP filter (left image). The transgenic insects fluoresce conspicuously compared to the wild-type under GFP. Notice how the recently eclosed beetle (the middle) is a lighter brown and also the ZsGreen is more conspicuous.

The transgenic males and females showed no fluorescence dimorphism under DsRed2 filter conditions indicating that the sex-alternate splicing of the transgene may not be functional.

5.3.5.3 MARKER RELIABILITY

The transformation marker was assessed in a similar fashion to the *Tc α-tubulin-1* and DsRed2 transgenesis module in OX4700 (5.3.1.3 Marker reliability). Again the reporter protein was regulated by *Tc α-tubulin-1* promoter, but with ZsGreen instead of DsRed2. The adult beetles

were scored for the fluorescent marker. The gDNA was then extracted for a PCR targeting the marker transgene (α -tubulin 5' UTR and ZsGreen) (Figure 113).

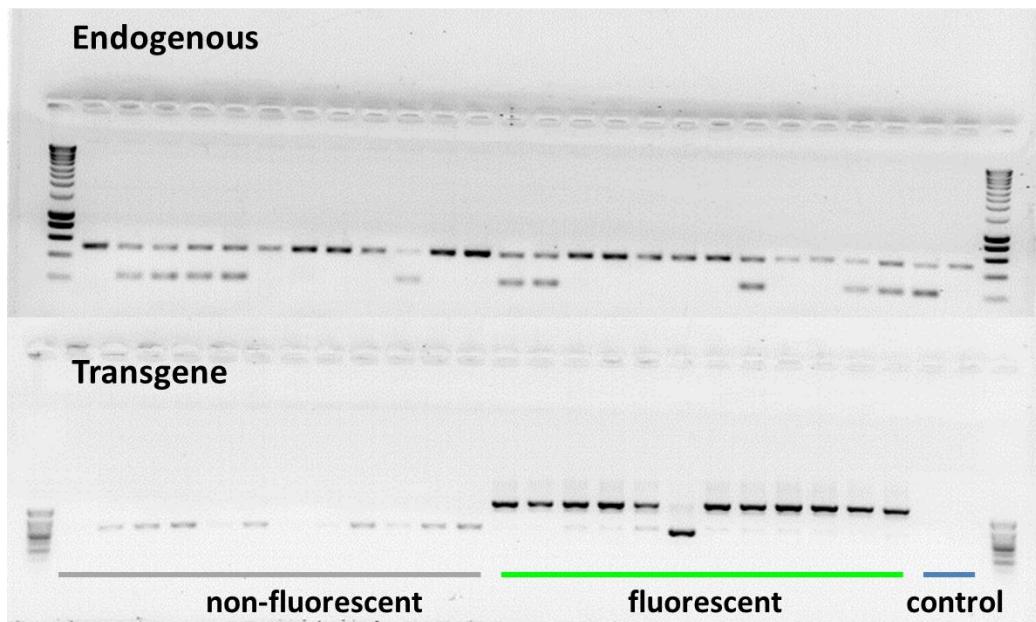


Figure 113. *α1tubulin-ZsGreen* in *Tribolium* is a reliable marker. A PCR was run using the extracted DNA of adult insects sorted by fluorescence. The endogenous control assessed the quality of the DNA and also provided a genetic means to assess the sex of the insects. Primers TD3372, 3373 and 3374 were used in a multiplex PCR with both sexes having a band at 500 bp and a male-specific band at 230 bp. Genomic possession of the marker transgene was assessed using primers across ZsGreen, TD3692, and the α -tubulin-1 promoter and 5' UTR, TD3026. The expected band size was 1130 bp. A negative control of a wild-type adult and OX4700Z was used both amplifying for the endogenous DNA but not the marker transgene. Surrounding lanes of the top row provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder). Surrounding lanes of the bottom row provide DNA size standards. Sizes are as follows: from bottom, 100 bp increasing in 100 bp increments until 800 bp, then 1000 bp (Eurogentec, SF Smartladder).

The marker's reliability was quantified by comparing the fluorescence status with PCR results.

The ZsGreen reduced the probability of mis-scoring a transgenic insect compared to OX4700C.

Only one misdiagnosis in 24 samples occurred (Table 32).

Table 32. A contingency table summarising the marker reliability experiment where the validity of ZsGreen screening (either fluorescent or non-fluorescent (wild-type)) for transgenics was evaluated against a gold standard of DNA extraction and transgene status confirmation using PCR (either transgene containing or non-transgene containing (wild-type)). The letters are provided to assist with term definition; where a) true positives, b) false positives, c) false negatives and d) true negatives. The test correctly identified only 11 of the 12 transgenic individuals giving a sensitivity of 100% (a / a+c). The test was worse on specificity (b / b+d) with one false positive. The positive predictive value (a / a+b) was 92%. This suggests on average a screener may miss one in ten transgenic adults. The negative predictive value was 100% (d / c+d). The overall accuracy was 96%. This gives a biased estimate as we have to correct for chance agreement. The chance level of agreement is 0.5. The corrected measure of agreement (kappa) is therefore: $\kappa = [\text{observed agreement} - \text{chance agreement}] / [1 - \text{chance agreement}] = [0.98 - 0.5] / [1 - 0.5] = 0.96$.

| Test result | True status | | Total |
|---------------|-------------|-----------|-----------|
| | transgene | non | |
| fluorescent | 11 (a) | 1 (b) | 12 |
| non | 0 (c) | 12 (d) | 12 |
| Totals | 11 | 13 | 24 |

With a Cohen's Kappa value of 0.96 even after correcting for a chance agreement we have almost a perfect level of agreement (near 1). This contrasts with the poor performance of DsRed2 (OX4700, $\kappa = 0.12$). It should be stressed that the test was conducted on just the adult insects and not the other life stage that is also typically screened (larva).

5.3.5.4 OX5118 DID NOT PRODUCE DSRED2 EXPRESSING FEMALES

No DsRed2 was detected by fluorescence microscopy (not shown) so a one-step RT-PCR was used to reveal if splicing was occurring as intended (using primers to amplify across *Tctra* and DsRed2). Transcript was present in both sexes with splicing of the shared intron and male exons occurring (Figure 114).

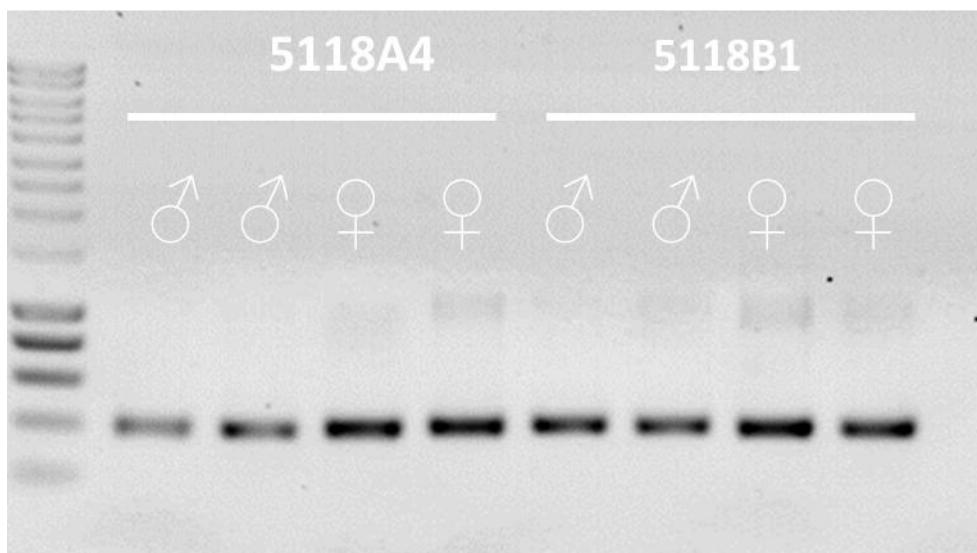


Figure 114. Using RT-PCR to assess sex-alternative splicing of OX5118 in transformed lines of *Tribolium*. The cDNA product of one-step RT-PCR amplified across the transgene (*Tctra* and DsRed2 using primers TD3026 and TD3692). Splicing of the shared intron and the male exon is occurring albeit with no alternate-splicing between the sexes (with a 400 bp shared intron and male-specific exons of 712 bp). For comparison the expected gdNA band size was 1130 bp. This does not show whether the splice variants are identical but shows they are equivalent in size. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

To check the DsRed2 was in-frame and to confirm the identity of the sequence after splicing, the RT-PCR product from one male and one female of each line was purified and sequenced. These sequences were first compared between lines to determine whether splicing was similar for each insertion.

The splicing occurred at the expected site, with DsRed2 in frame (Figure 115). However, the sequenced OX5118A male amplicon did not contain the male exon suggesting it had been spliced from the transgene mRNA. Exon-intron boundaries were confirmed by aligning the sequences.

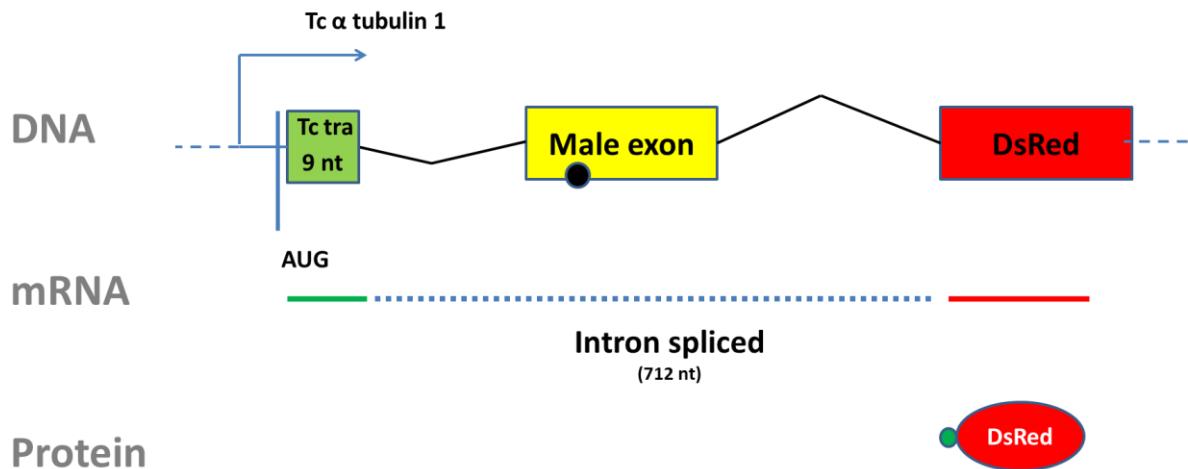


Figure 115. Splicing of the OX5118A and B transgene transcript in *Tribolium castaneum*. In both sexes the entire 712 bp intron is spliced out, in contrast to the endogenous *transformer* where the male exon is not spliced out in males, causing non-functional truncated protein due to a stop codon.

The same work flow was applied to the females of OX5118A and B, in which the same conclusion was made: both lines spliced in the same way. The inter-sex comparison also showed no difference in splicing, confirming the PCR result of no sex-alternate splicing (Figure 114).

Critical to the designed functionality of OX5118 was the alternate splicing between the sexes with the non-functional mRNA (exon with stop codon) produced in the male and functional mRNA in the female. However, the lack of detectable DsRed2 fluorescence in either sex was unexpected as the sequence was in frame and correct (albeit with a couple of amino acids extra on the 5' end as per the intended design).

Double-checking the intronic sequence of OX5118 compared to Shukla & Palli *Tctra* sequence (2012b) revealed it had been mislabelled; the male exon had not been included in the design. This revelation resulted in a re-design by including the originally intended male-specific exon as well as a small number of nucleotides either side of the splice site.

5.3.5.5 OX5118 DSRED2 EXPRESSION RE-EXAMINED

Previously when screening for DsRed2 we expected a difference between sexes, females positive and males. Perhaps both sexes were positive for DsRed2, but it was not detected due to the lack of a negative control?

RT-PCR revealed that both sexes expressed DsRed2 in frame. I examined and compared OX5118 against wild-type (negative for DsRed2) and OX4700 (positive for DsRed2). No DsRed2 protein was detected in OX5118 (Figure 116).

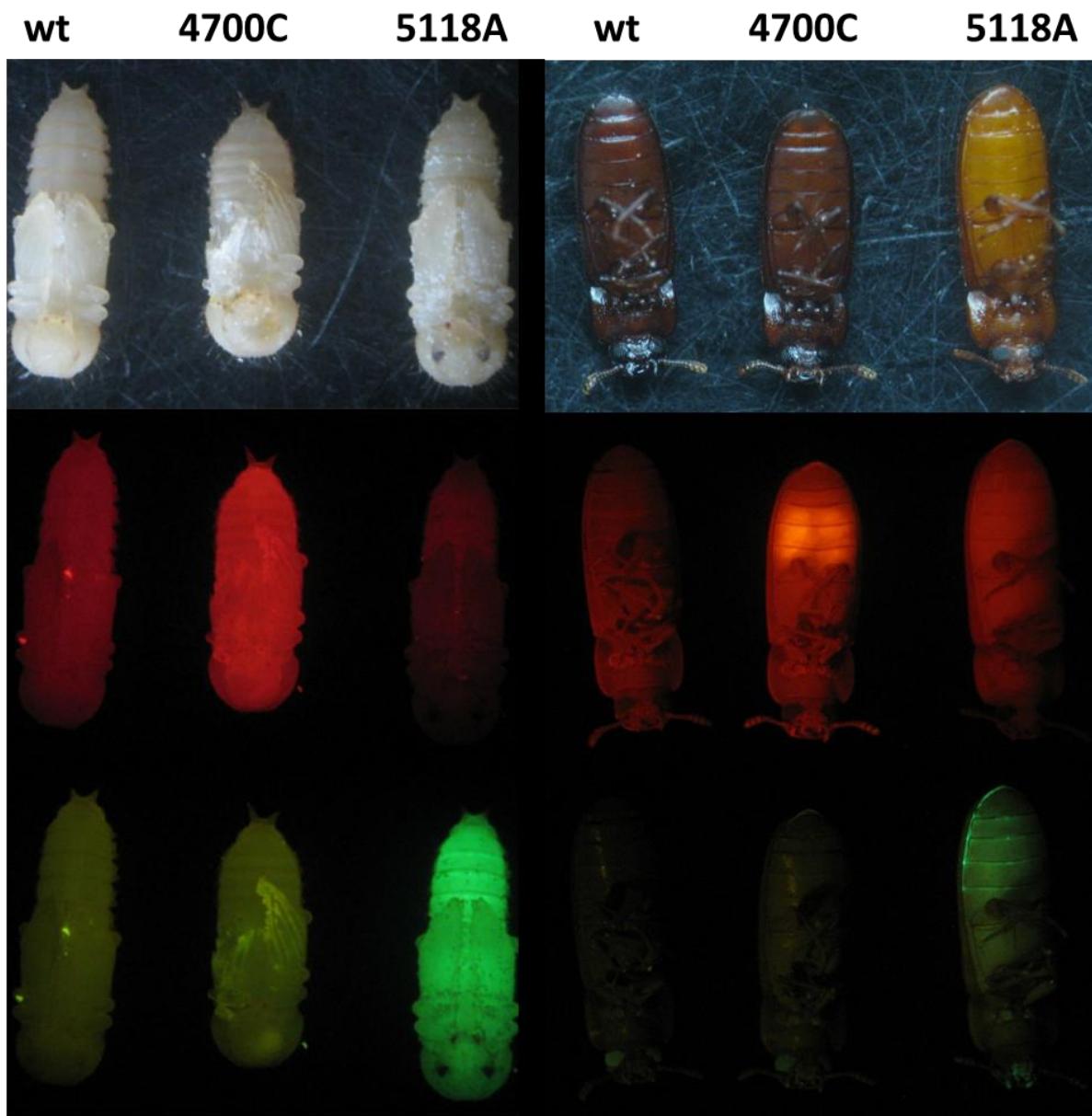


Figure 116. Comparison of male adults and pupae of different *Tribolium castaneum* lines with wild-type counterparts (wt) using fluorescence microscopy; under white-light (top), DsRed2 (middle) and GFP filters (bottom) and lighting conditions. The wild-type is negative for DsRed2 and ZsGreen, OX4700C is negative for ZsGreen but positive for DsRed2 and OX5118A is negative for DsRed2 but positive for ZsGreen. Larvae are not shown but exhibited similar marker expression. Males were used as the fluorescence in the testes as adults is conspicuous.

The failure to detect the DsRed2 protein was hard to explain. It may have been due to three amino acids of the common exon from *transformer* next to the intron being included and interfering with DsRed2 formation and tetramisation (Figure 115).

It was not determined why, in OX5118A and B, DsRed2 was not being translated to a detectable level using fluorescence microscopy. Antibodies could have been used to detect DsRed2, but this

was deemed lower priority relative to developing an alternatively spliced sexing system. Lack of visible DsRed could be due to the weakness of the Kozak consensus sequence being used in the construct. The promoter sequence was re-inspected to determine any differences between upstream DNA of the fluorescent protein coding sequences, comparing the transgenesis module with the female-specific module.

5.3.5.8 DSRED2 FAILED TRANSLATION IN OX5118 MALES AND FEMALES

ZsGreen was visibly expressed in OX5118 whereas DsRed2 was not. Both shared identical promoter regions based on the whole 5' UTR of α -tubulin-1. The only difference was the inclusion of a putative intron upstream of the ZsGreen. If this intron is not removed during splicing then it may serve a role during translation, explaining the discrepancy between the different fluorescent proteins translation status. A one-step RT-PCR was conducted to determine if this was the case (Figure 117).

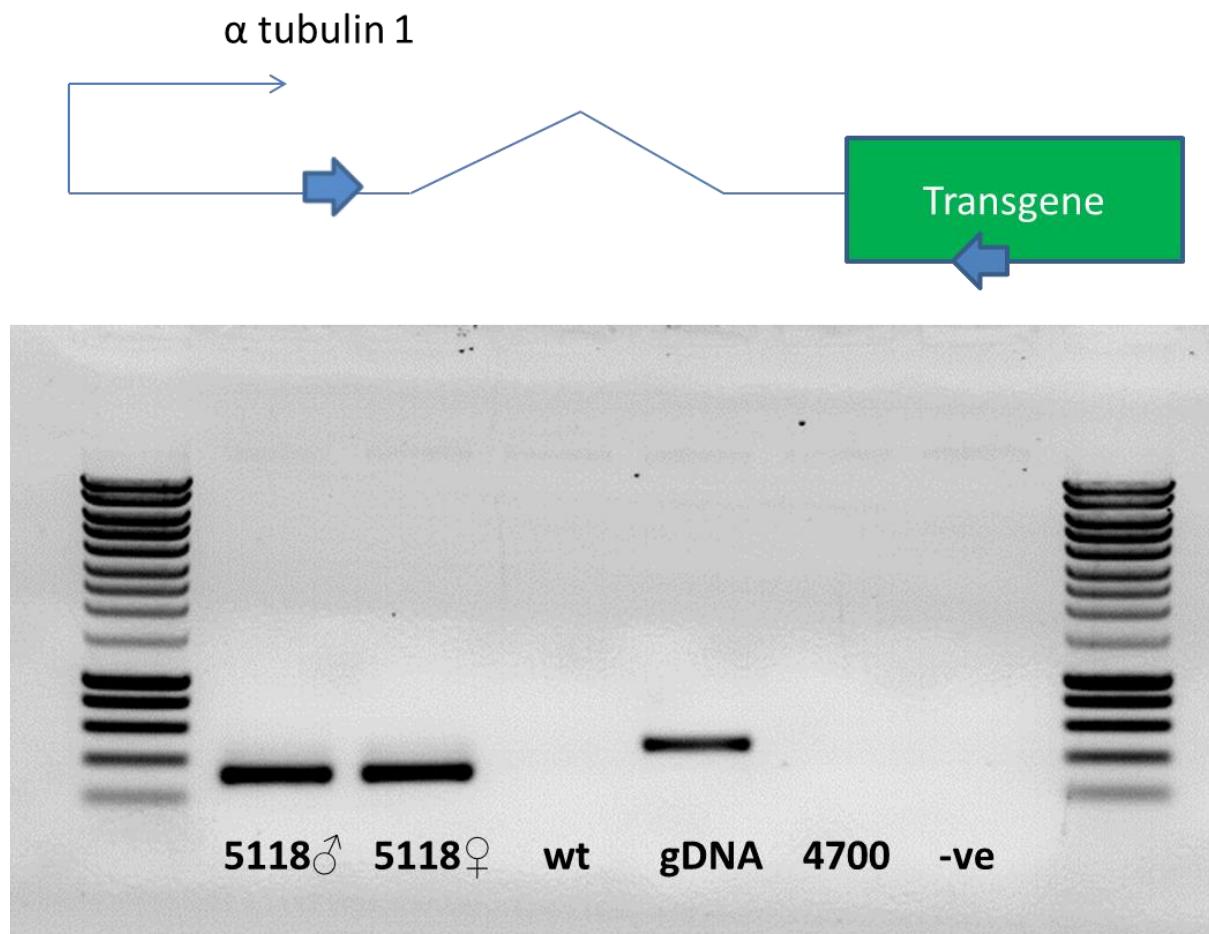


Figure 117. Assessing whether the intron (wavy line) between the promoter (thin arrow) to the 5' UTR (between the intron and the transgene) of the marker transgene *α -tubulin-1-ZsGreen* is spliced prior to translation in OX5118A. There is discrepancy in size between the gDNA amplicon and the OX5118 cDNA amplicon equivalent to the size of the intron. The target sequence for amplification in OX5118 is represented diagrammatically (top). The one-step RT-PCR was conducted using the forward primer TD3026 and reverse primer TD802 targeting the main marker transgene *α -tubulin-1-ZsGreen*. The gel shows the cDNA product of one-step RT-PCR amplifying across the *α -tubulin-1* endogenous promoter, 5' UTR and intron. Splicing is occurring with no differences between the sexes. For comparison the expected cDNA band size was 316 bp. The intron was 154 bp in length and was successfully spliced. No product was produced for those sequences without ZsGreen. OX5118 both sexes were positive for the transgene 5' UTR with the intron removed. Wild-type gDNA (wt), OX4700 gDNA and a control of water were negative for the transgene. The OX5118 gDNA formed a band at 470 bp. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The gel suggested that the intron was removed from the transcript of the OX5118 marker *α -tubulin-1-ZsGreen*. This data implied the omission of the intron from the OX5118 *Tctra-DsRed2* transgene was not responsible for the failed translation. Failing to understand this problem, in designing the next construct OX5145, identical promoter and non-coding sequence were used upstream of the alternatively-splicing sexing marker as was used in the design of the transformation marker.

5.3.6 OX5145 - WHAT OX5118 WAS INTENDED TO BE!

OX5145 is similar in design to that of OX5118, but accurately included all of the male exons and shared intron (Figure 109). The premise of the new design was that successful splicing in the females would produce functional DsRed2 in females and non-functional DsRed2. The promoter region for the DsRed2 was matched to that which had worked previously in OX4700 albeit with the addition of the *Tctra* minigene.

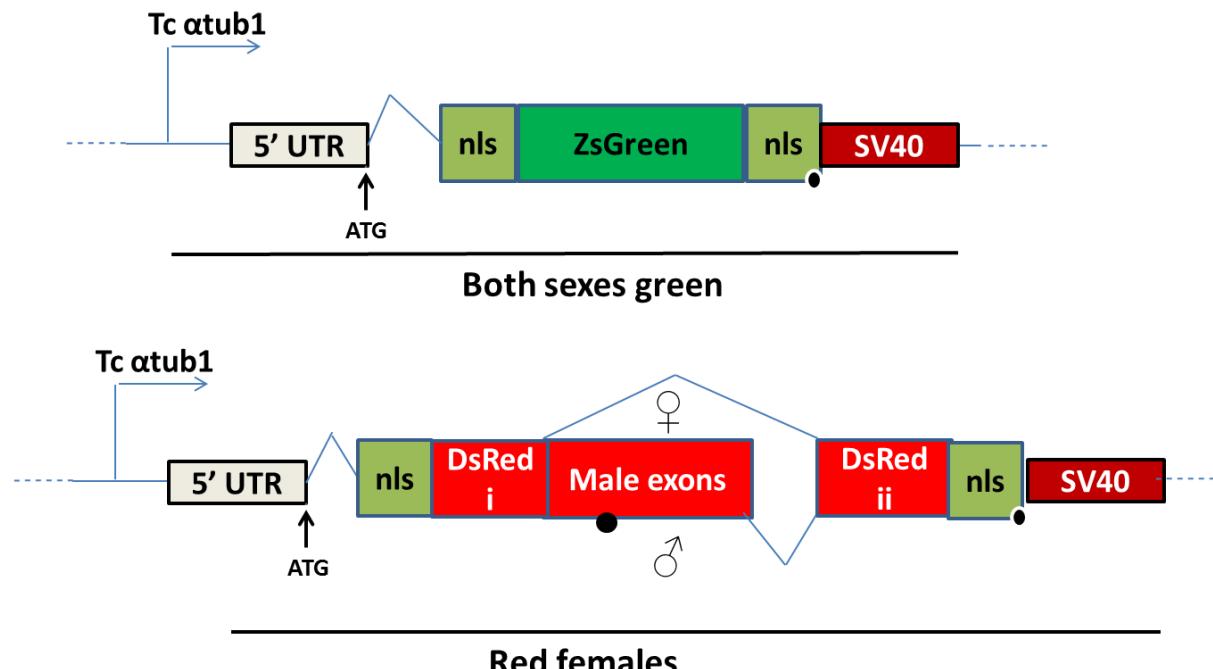


Figure 118. The improved design of the genetic sexing construct in *Tribolium castaneum*, OX5145. The transformation marker is shown on the top. The female-specific component was dependent on alternative splicing between the sexes (bottom). The females would splice out the male exons and stop codon therein, producing functional DsRed2 protein. Stop codons are shown by black circles.

5.3.6.1 INJECTION DATA

The number of embryos injected (Table 33) was based on logistics and advice from the “Goldilocks” model (Chapter 2). The transformation efficiency was higher for OX5145 (~ 8%) compared to OX5118 (~ 1%).

Table 33. Injection data for OX5145 injected into *Tribolium* embryos. G₀ are the number of fertile adults derived from the injected embryos. G₁ were not counted to save time. The number of independent transgenic lines, X (conservative – only those from unique crosses were considered

independent lines). Notice the eight fold increase in transformation efficiency using ZsGreen compared to DsRed2.

| Injected embryos | Larvae | G₀ | G₁ | X |
|-----------------------------|---------------|----------------------|----------------------|----------|
| 918 | - | 108 (0.12) | - | 9 (0.08) |

On separating the ZsGreen positives the insects were exposed to DsRed2 screening conditions but no convincing fluorescence was detected at larval, pupal (Figure 40) and adult (Figure 45) life stages when compared to wild-type counterparts.

5.3.6.2 TRANSGENESIS

Transformants were detected when screening injection survivor offspring for ZsGreen expression. Although ZsGreen becomes more difficult to detect as the pupae mature and the elytra darken, the marker was far superior to DsRed2 based on marker reliability assays and my opinion after screening tens of thousands of beetles.

The ZsGreen was expressed in both sexes with the DsRed2 not detectable in either sex at the pupal stage (Figure 119).

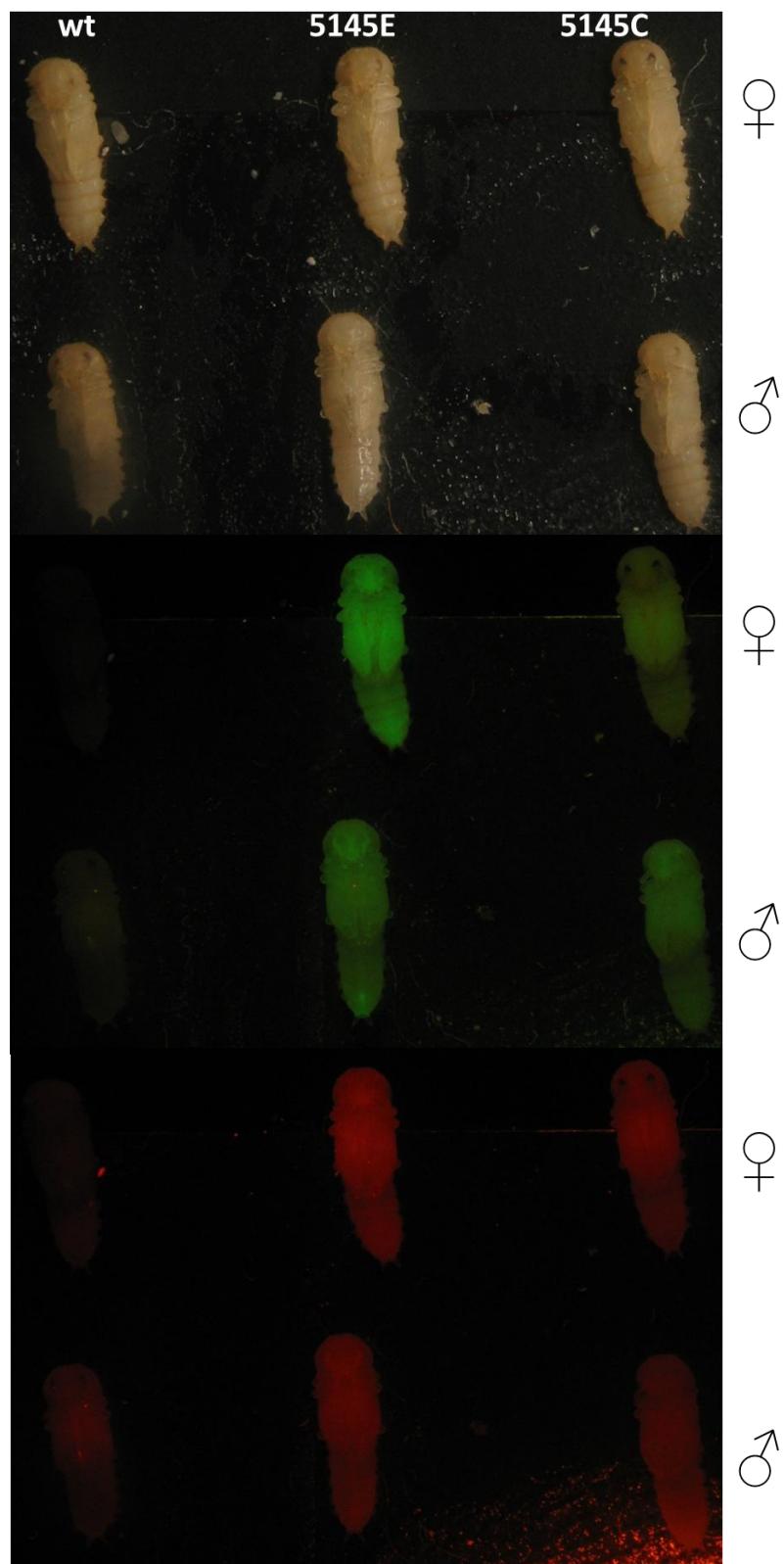


Figure 119. OX5145 pupae assessed for ZsGreen and DsRed2 using fluorescence microscopy. The putative transgenic pupae (middle and right columns) are compared to wild-type pupae (wt; left column) with both sexes shown under three lighting conditions (from top to bottom; white light, GFP, DsRed2). Both sexes were positive for GFP with DsRed2 status uncertain. This does not necessarily mean DsRed2 is absent as this promoter marker combination has been shown to be poor

especially at the pupal stage. The image could be improved with a DsRed2 positive control of OX4700. The zoomed out nature shows typical screening magnification used. Notice how trapped flour can cause auto-fluorescence.

Transformation events were confirmed by PCR (Figure 120A).

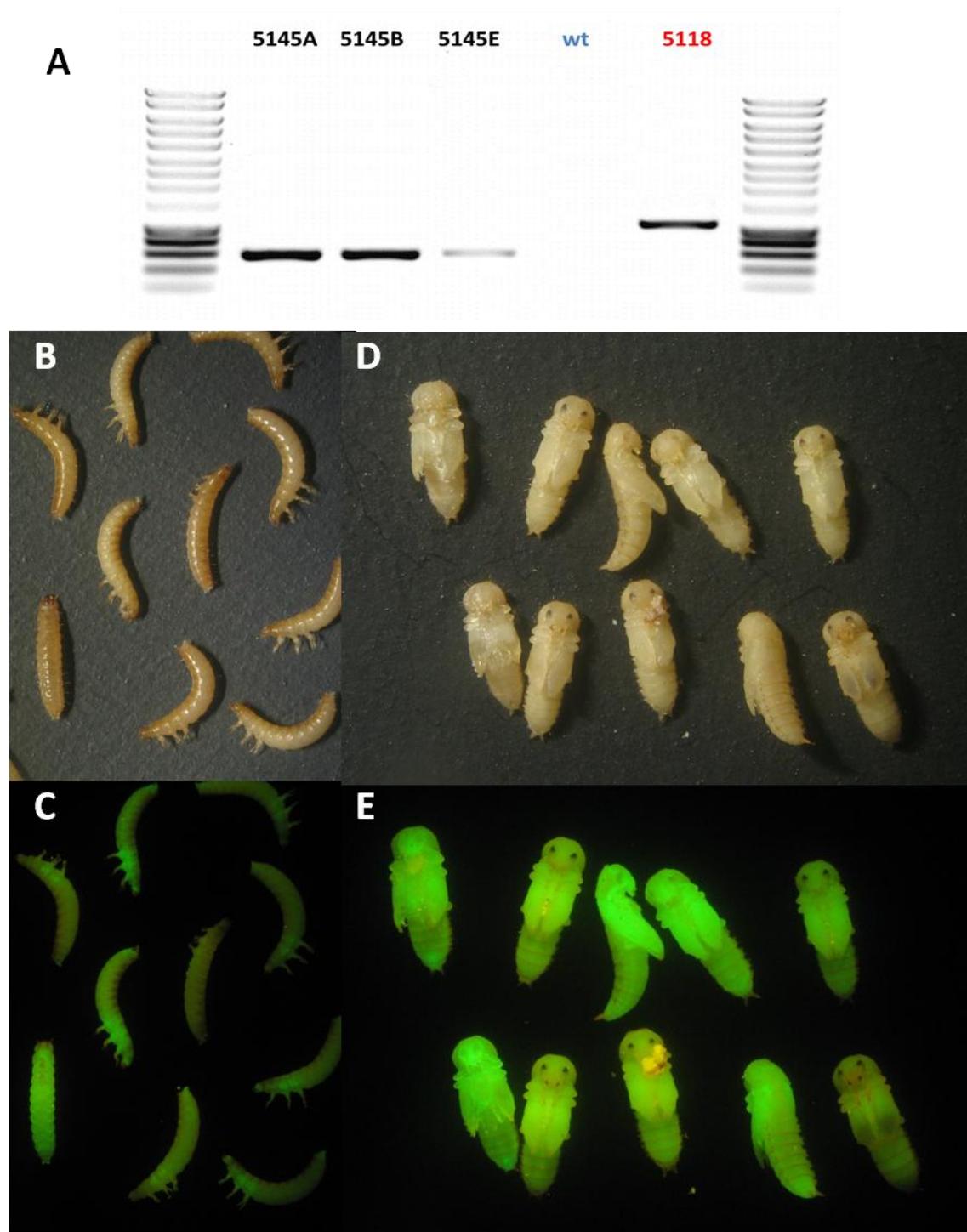


Figure 120. OX5145 transgenesis confirmed using PCR and fluorescence microscopy. A) PCR confirmation of OX5145 transformants. The PCR amplicons from a reaction involving different gDNA samples were run on a gel. The expected band size of the PCR products of gDNA with primers TD3026 and TD3692 was 609 bp. The same reaction was expected to produce a band size of 1100 bp with gDNA from OX5118 (the positive control) and no band for the gDNA of wild-type (wt, as the negative control). Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder). B & C) ZsGreen expression detected using fluorescence microscopy in

OX5145 larvae driven by the *Tca-tubulin-1* promoter. Expression is located mainly in the elytra discs of the larvae. Notice the phenotypic variation which may be associated with the darkening of the larvae as they mature (ZsGreen filters C, bright field B). D & E) The ZsGreen driven by the *Tca-tubulin-1* promoter is located mainly in the elytra of the pupae. Notice how the more mature and melanised elytra of the older pupae (blackened elytra) reduce the amount of light escaping the insect (ZsGreen filters E, bright field D)

In order to perpetuate the line and assess expected Mendelian inheritance based on one insertion event per line, crosses were set up with wild-type *Tribolium*. All lines were identified by ZsGreen, no lines showed any evidence of DsRed2 expression for both sexes.

Lines were assessed for multiple insertions and to provide sufficient numbers to confirm no DsRed2 expression in either sex (Table 34).

Table 34. Assessing the OX5145 lines for multiple insertions by crossing G₁ with wild-type. Putative transgenic G₁ insects derived from unique G₀ crosses were separated and labelled as different putative lines. One transgenic male was crossed with five wild-type females (or one female with three wild-type males if no females were produced by a particular G₀ cross). The frequency of transgenic to non-transgenic insects after three weeks were compared to gauge the likelihood of multiple transgene insertions in a line indicated by non-Mendelian ratios (or a proportion significantly greater than 50% transgenic (indicated by EGFP fluorescent protein marker) to non-transgenic phenotype). Evidence for multiple insertions of the transgene in any line was weak except possibly in line H. Two tailed chi-square with one degree of freedom. Proportions given in brackets.

| Line | Transgenic | Non-transgenic | X ² (2 d.p.) | P (2 d.p.) |
|----------|------------|----------------|-------------------------|------------|
| OX5145A1 | 42 (0.45) | 50 (0.55) | 0.70 | 0.40 |
| OX5145A2 | 22 (0.65) | 12 (0.35) | 2.94 | 0.09 |
| OX5145B1 | 40 (0.35) | 76 (0.65) | 11.1 | <0.01 |
| OX5145B2 | 50 (0.53) | 45 (0.47) | 0.26 | 0.61 |
| OX5145C1 | 42 (0.49) | 44 (0.51) | 0.05 | 0.83 |
| OX5145E1 | 73 (0.53) | 65 (0.47) | 0.46 | 0.50 |
| OX5145H1 | 46 (0.65) | 25 (0.35) | 6.21 | 0.01 |

Line OX5145B1 was examined again with an additional wild-type cross and no significant difference was found. Given the number of chi-squared tests conducted, type one errors become more likely (incorrect rejection of H₀). Lines were kept for several generations through testing – showing inheritance of the transgene for >2 generations - then all but one discarded.

The elytral discs in male and female pupae and larvae were examined for both ZsGreen and DsRed2. No DsRed2 was visible whereas the ZsGreen was visible in both sexes with an expression pattern similar to that of the transformation marker used in OX4700 and OX5118

(Figure 120). Variation is evident in both the larvae and the pupae although as the pupae mature and the elytra darken the ZsGreen becomes less conspicuous (Figure 120). These images support the use of ZsGreen as the main marker compared to DsRed2, and demonstrate the phenotypic variation of this transformation marker.

Microphotograph of the testes also corroborated the findings of the expression pattern concentrated in the testes as adults (Figure 121), confirming observations of Siebert *et al.*, 2008 and in OX4700C, described earlier in this Chapter (5.3.1.4.1 Male).

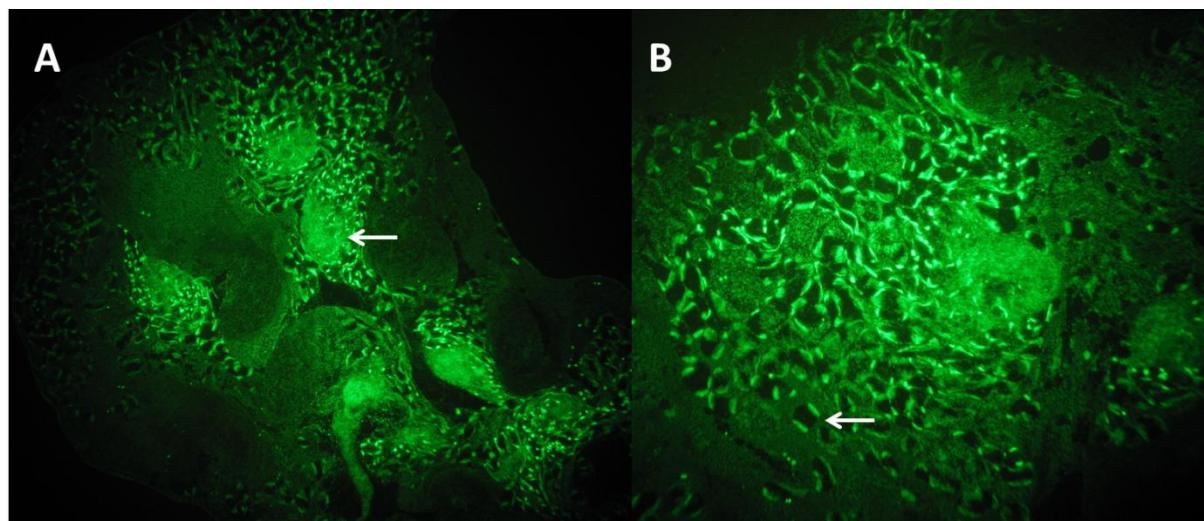


Figure 121. Confirmation of the expression profile of *Tc- α -tubulin-1-ZsGreen* in adult testes. A) OX5145 testis under a cover slip at x40 magnification. The arrow shows one follicle where the high density of cells undergoing spermatogenesis cause bright spot relative to the surrounding darker somatic tissue. B) At x200 magnification individual cysts can be seen with putative sperm bundles drawing to either pole of the cyst.

5.3.6.3 FEMALE EXPRESSION COMPARISON

Examination of OX5145 female pupae for DsRed2 expression was inconclusive. Following removal of elytra and internal examination of the female reproductive organs, DsRed2 expression was assessed by comparing with negative (wild-type) and positive (OX4700C) controls using fluorescence microscopy (Figure 122A). Females were starved for a week to remove the opaque, assumed to be fat, layer for inspection by fluorescence microscopy (Figure 122B). Female-specific DsRed2 expression in the ovaries did not mirror that of the ZsGreen. Only auto-fluorescence of the gut and or its contents was detected. However, given the insects were starved this could be DsRed2 expression.

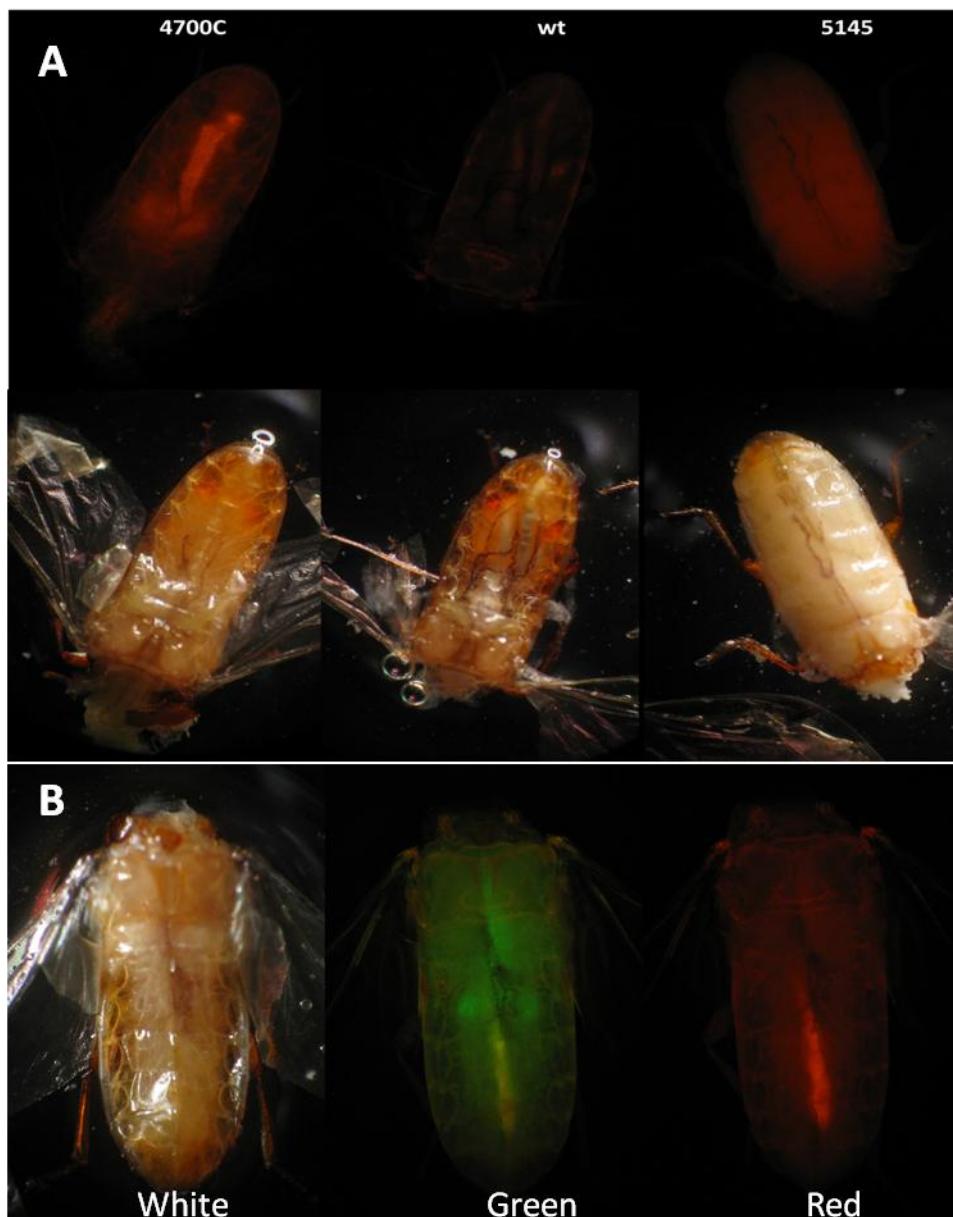


Figure 122. OX5145 females were dissected and examined for DsRed2 expression using fluorescence microscopy. (A) Forty times magnification of three randomly sampled insects from the lines OX4700C, wild-type (wt) and OX5145. In adulthood the transgene is expressed in specific parts of the female presumably the reproductive system in a pair of spherical organs in the lower abdomen. This fluorescence is not detectable in wild-type nor OX5145 insects (possibly due to opaque fat deposits obscuring the view). The *a-tubulin-1* promoter appears to drive expression in the ovaries. (B) Starved OX5145 females were dissected and examined for DsRed2 expression using fluorescence microscopy No DsRed2 was detected in OX5145 females except for auto-fluorescence of the gut or its contents. Forty times magnification with additional camera zoom detailing the fluorescence patterning of OX5145 female. The head, thorax and elytra have been removed and the fatty layer lost due to starvation permitting visual inspection through the dorsal side of the insect. The gut passes close to the reproductive system and it is the flour contained therein which is thought to give rise to the auto-fluorescence therein. The insects were starved for a week and then randomly sampled. The ovaries can be seen to be positive for ZsGreen (Green) but not for DsRed (right).

5.3.6.4 COMPARING TRANSFORMATION MARKER TRANSCRIPT IN MALES AND

FEMALES

Correct splicing in OX5145 would have resulted in a female-specific splice variant which would be translated into functional protein and two splice variants in males, both coding for non-functional protein due to the stop codon in male exons 2a and 2b. Splicing of the transcript was assessed by RNA extraction from males and females from two different lines followed by a one-step RT-PCR using primers targeting DsRed2. The band sizes from the females were of the expected size implying shared intron and male exon removal (Figure 123).

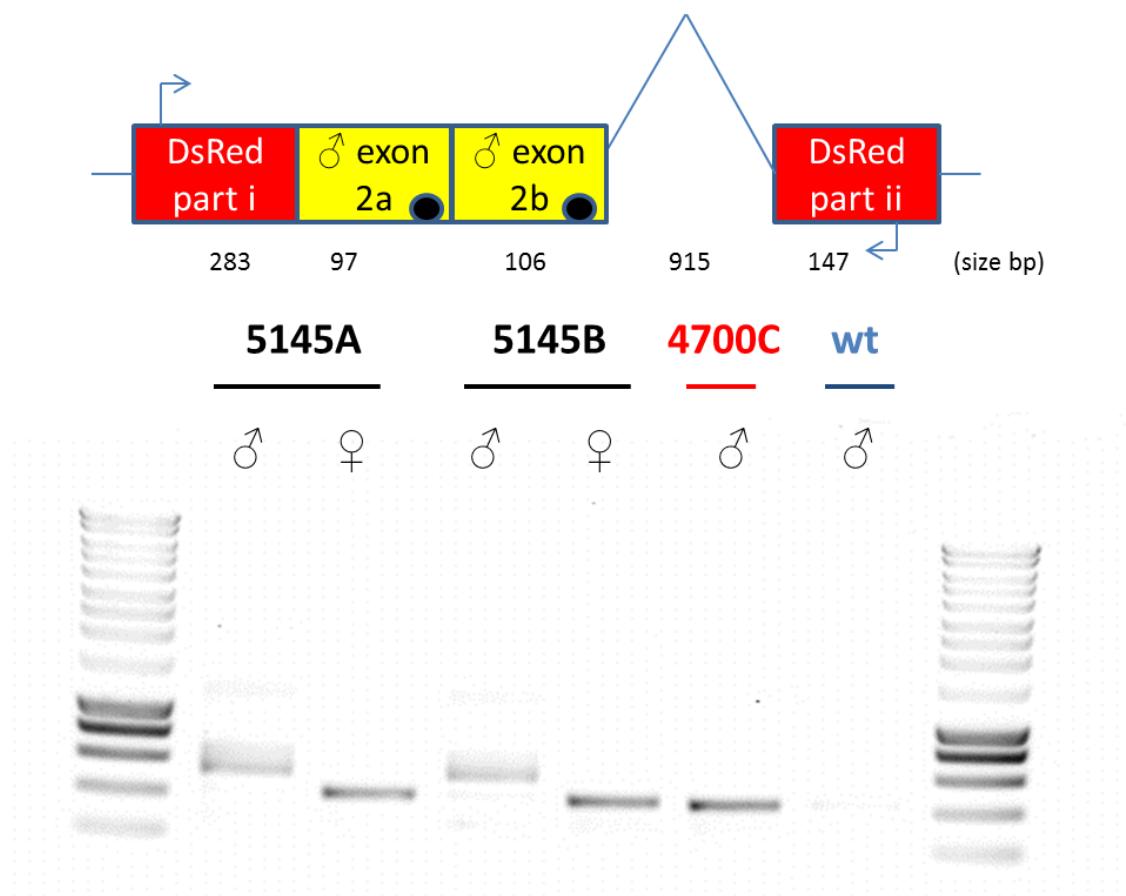


Figure 123. Alternative splicing of the *Tctra* minigene in OX5145 insects. A diagram of the transgene design in OX5145 where the female transcript is smaller than the male due to alternative splicing (top). A gel of the RT-PCR products showing the transgene associated transcript with primers (either side of the splice site; TD94 and TD95) targeting DsRed2 and the *Tctra* intron therein (bottom). RNA was extracted from a male and female pupa of two lines A and B. No splicing would result in a band size equivalent to gDNA of 1345 bp. Expected splicing in the female would remove the shared intron (915 bp) and the two male exons giving a band of 430 bp. The male exons contain stop codons represented by black filled circles. The male is expected to produce two isoforms of the transcript with either just 2a included (536 bp) or 2a and 2b included (633 bp). Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The amplicons from the females were gel-purified, cloned and sequenced. The reverse complement of each sequence read was taken and compared to the hypothetical transcript of a successfully spliced female with all introns removed (Figure 124).

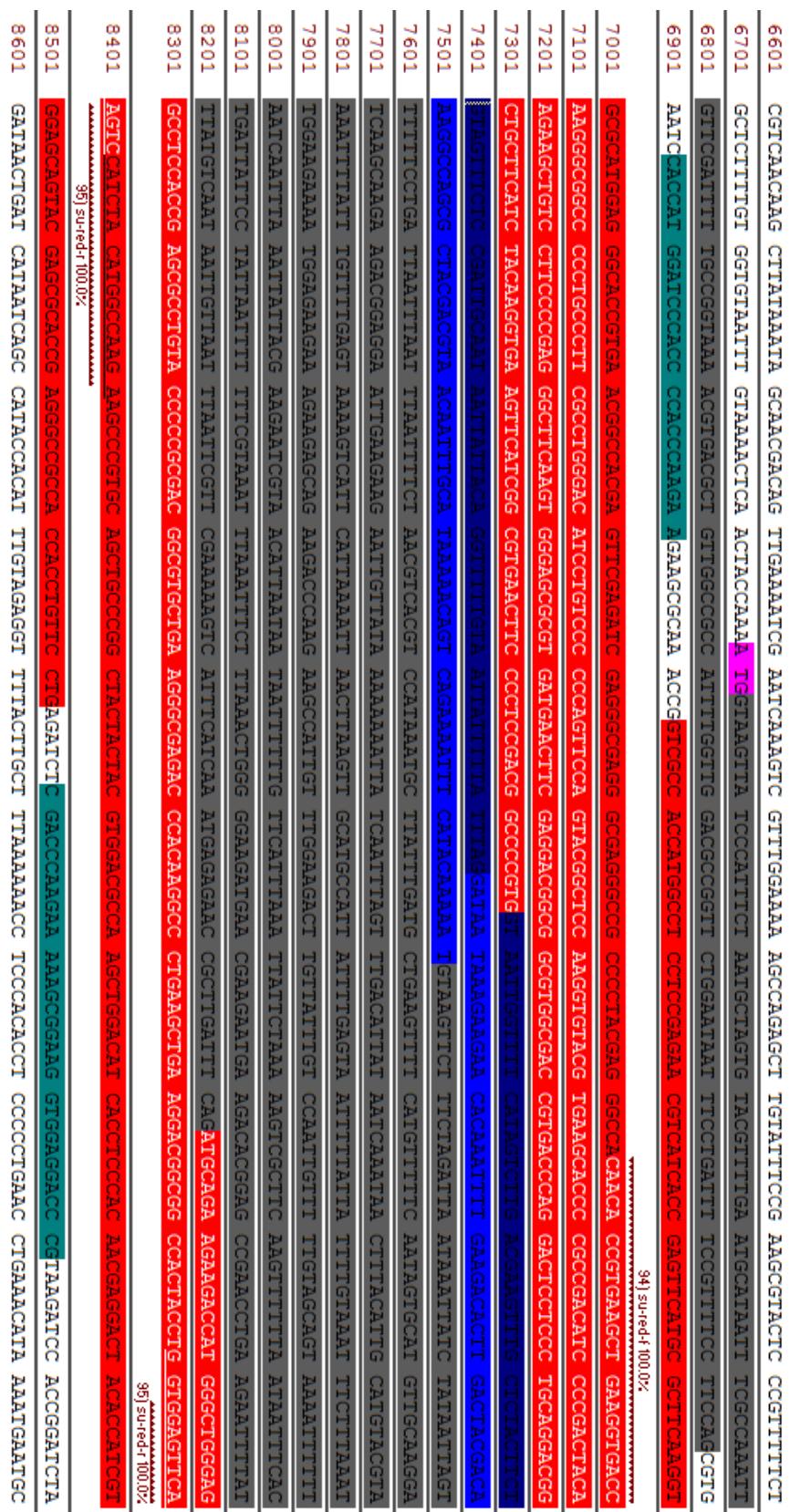


Figure 124. Splicing of the *tra* minigene in OX5145 females matched expectation. The transgene region of OX5145 designed to produce functional DsRed2 (due to splicing of stop codon containing male exons) in females. The females of OX5145A successfully spliced out both male exons and the shared intron to produce DsRed2 transcript, however the 5' and 3' ends were not covered by the sequencing (the primers were chosen to span the introns to resolve whether splicing was occurring

as anticipated). The transcript sequence is highlighted on the OX5145 as the white font of the nucleotides, matching the design perfectly. The transcript for OX5145B was probably non-specific as it only matched sequence for the reverse primer in the second part of DsRed2 (35 bp), shown by underline. Only the 5' strand of the sequence is shown for clarity. The putative start codon is highlighted pink (note also the ATG in the cyan highlighted nuclear localisation signal region and near the start of the red highlighted DsRed2 first exon, both have Kozak sequences). The male-specific exons 2a and 2b are dark blue and light blue respectively, both are spliced out by the female. Shared introns are shown by the grey region with the 5' UTR a darker shade of grey. The amplified region of the one-step RT-PCR using the primers TD 94 and 95 is shown, with the primer regions shown by the dark red serrated line.

The promoter region, 5' UTR and start site of the DsRed2 region of OX5145 were identical to our positive control OX4700C. The only difference was the addition of the *Tctra* minigene, which was shown to be successfully spliced out in females. One would expect DsRed2 translation given this information as shown with the addition of the amino acid sequence to the figure (Figure 125). The sequence appears to be in-frame and the DsRed2 chromophore intact.

| | | |
|------|---|---------------------|
| | | Met |
| 6701 | GCTCTTTGT GGTGTAATT GTAAAAACTCA ACTACCAAAA TGTTAAGTT TCCCATTCTT AATGCTAGTG TAGCTTTGA ATGCATAATT TCGCCAAATT | Arg Glu |
| 6801 | GGTCGATT TTGCGCGTAA ACGTGACGCT GTTGGCCGCC ATTGGGTTT GACGCCGTT CTGGAATAAT TTCTGATT TCCGTTTCC TTCCAGCGT | |
| 6901 | Glu Ser Thr Met Asp Pro Thr Pro Pro Lys Lys Lys Arg Lys Pro Val Ala Thr Met Ala Ser Ser Glu Asn Val Ile Thr Glu Phe Met Arg Phe Lys Val | |
| | AATC CACCAT GGATCCCACC CCACCCAAGA AGAACGCAA ACCGCTGCC ACCATGGCT CCTCCGAGAA CGICATCAC GAGITCAIGC GCITCAAGG | 94) su-red-r 100.0% |
| +3 | Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val Thr | |
| 7001 | GCGCAIAGG GGCACCGTGA ACGGCCACGA CTICGAGATC GAGGGCGAGG GCGAGGGCG CCCTIACGAG GGGCACRACA CGTGAAGCT GAGGAGGACCG | |
| +3 | Lys Gly Gly Pro Leu Pro Phe Ala Ttp Asp Ile Leu Ser Pro Gln Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro Asp Tyr Lys | |
| 7101 | AGGGCGGCC CCCTGCCCTT CGCTGGGAC ATCCGTGCCC CCCAGTTCA GTACGGGCC AAGGTGTAAG TGAAGCACCC CGCCGACATC CGGCACTACA | |
| +3 | Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Ala Thr Val Glu Asp Ser Ser Leu Glu Asp Gly | |
| 7201 | AGAAGCTGTC CTCCCCGAG GCTTCAAGT GGGACCGCGT CATGAACTTC GAGGACGCG GCGTGGCGAC CGTGAACCCAG GACTCCCTCC TGCAAGGACCG | |
| +3 | Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Glu Val Asn Phe Pro Ser Asp Gly Pro Val | |
| 7301 | CTGCTTCATC TACAAGGTGA AGTCAIACCG CGTGAACCTC CCCTCCGAGC GCGCGGTGT AATIGGTTT CATAIGCTTG ACCAGGTTT CTCTACTCTT | |
| 7401 | TGATGTTCTC CGATTGCAAT ATTATTACA GGTTTTGTAA ATTATTTTTA TTAGGATAA TAAAGAAGAA CACAAATTTC GAAGACATT GACTACGACA | |
| 7501 | AAGGCCAGCG CTACGACGTA ACAATTGCA TAAACAGT CAGAAATTCT CATACAAAAAA TGTAAAGTTCT TTCTAGATTA ATAATTATTAC TATAATTAGT | |
| 7601 | TTTCTCTGA TTAAATTATA TTAAATTCTT AACGTCACGT CCATAATGC TTATTGATG CTGAGTTTTC CATGTTTTC AATAGTCAT GTGCGAAGGA | |
| 7701 | TCAAGCAAGA AGACGGAGGA ATTAAAGAAG AATTTTATA AAAAATTAA TCAATTAGT TTGACATTAT AATCAAAATAA CTTACATTG CTAGTCAGTA | |
| 7801 | AAATTATTATG TTGTTTGTAGT AAAAGTCATT CATTAAATT AACTTAAGTT GCATGCCATT ATTGAGATA ATTTTTATA TTGTTAAAT TTCTTTAAAT | |
| 7901 | TGGAAGAAAA TGGAGAAGAA AGAAAGACGAG AAGACCCAAAG AAGCCATTGT TTGGAAGACT TTGTTATTGT CCAATTGTTT TTGAGCAGT AAAATTTTT | |
| 8001 | AATCAATTAA ATTATTACG AAGAATCGTA ACATTAATAA TAATTTTTG TTCAATTAAA TTATTCTAA AAGTCGCTTC AAGTTTTTA ATAATTTCAC | |
| 8101 | TGATTATTCC TATTAATTTT TTGCTTAAT TAAACTGGG GGAAGATGAA CGARGAATGA AGACACGGAG CGAACCTGA AGAATTTTAT | |
| +3 | Met Gln Lys Lys Thr Met Gly Trp Glu | |
| 8201 | TTATGTCAT ATTGTTAAAT TTAACTCGT CGAAAAAGTC ATTTCATCAA ATGAGAGAAC CGCTTGATT CAGATGCGA AGAAGACCAT GGCCTGGGAG | |
| +3 | Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu Thr His Lys Ala Leu Lys Leu Lys Asp Gly Glu His Tyr Leu Val Phe Pro Lys | |
| 8301 | GCCTCCACCG ACCGCCGTGA CCCCGCGAC CGCGTGTGA AGGGCGACAC CCACAAAGGCC CTGAGCTGA AGGACGGCGG CCACCTACCTG TGCGAGCTCA | 95) su-red-r 100.0% |
| +3 | Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Glu Leu Pro Gly Tyr Tyr Tyr Val Asp Ala Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile Val | |
| 8401 | AGTCCATCTA CATGGCCAAAG AAGCCCGTGC AGCTGCCCGT CTACTACTAC TTGAGACGCC AGCTGGACAT CACCTCCAC AACGAGGACT ACACCATCGT | |
| +3 | Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe Leu Arg Ser Arg Pro Lys Lys Lys Arg Lys Val Glu Asp Pro ... | |
| 8501 | GGACGAGCTA GAGGCCACCG AGGGCGGCCA CCACCTGTTC CTAGAACCTC GACCCAAAGAA AAACGCGAAG GTGAGGACCC TAAAGATCC ACCGGATCTA | |

Figure 125. Splicing of the *tra* minigene in OX5145 females. The transgene region of OX5145 designed to produce female-specific DsRed2 (due to splicing of stop codon containing male exons). Only the 5' strand of the sequence is shown for clarity. The putative start codon is highlighted purple. The amplified region of the one-step RT-PCR using the primers TD 94 and 95 is shown, with the primer regions shown by the dark red serrated line. The expected amino acid sequence is shown and with the DsRed2 in-frame and intact with the chromophore shown (turquoise). The splicing was shown to be perfect in the sequenced RT-PCR female band data.

Perhaps access to the translational machinery of the cell is outcompeted for by the endogenous α -actin and the ZsGreen transgene transcript preventing the DsRed2 variant from access to translation. As not all the DsRed2 transcript was sequenced there is the possibility of an error towards the start codon. This is unlikely as it was “copied and pasted” from OX4700. It is unclear why the protein was not visible as the transcript is sequentially identical to that produced in OX4700 as the transformation marker.

5.3.6.5 HOW TO PROCEED GIVEN TRANSCRIPTION AND ALTERNATIVE SPLICING BUT NO TRANSLATION

Given confirmation of the transgene transcript sequence, that we attempt conditional female-specific transcription and lethality.

I proposed a conditional female-lethal design (Figure 126). This was preferred to a reporter line system due to a limitation in the number of promoter/reporter combinations available. If the tetracycline in the diet were not being taken up by the females then at least I would detect skewed sex ratios by crossing male transformants with wild-type females. DsRed2 coding sequence was included as another reporter due to the bidirectional nature of *tetO*. The α -*tubulin-1* promoter upstream of the tTAV coding sequence could have been replaced with polyubiquitin or *hsp70*, but the endogenous promoter was preferred as it had been tested previously. The design was intended to provide conditional lethality to females, regulated by tetracycline in the diet.

Bisecting the tTAV coding sequence with the *tra* intron provides an alternative (to the design of OX5118). I opted for this strategy as I could not find appropriate sequence to match the endogenous splicing requirements. I also considered removing the endogenous intron and exon sequence following the *Tc* α -*tubulin-1* promoter as tTAV protein is sensitive to amino acids on its N terminal. I chose not to as in some genes of *Tribolium* and *Drosophila* it has been shown that sequences within the first intron are required for transcription (O' Donnel, 1994; Siebert *et al.*, 2009).

5.3.7 THE LAST HURRAH

Based on these assumptions the next construct was designed. Tarig Dafa'alla assisted with the design, suggesting that tTAV needs to be separated from any additional amino acids by ubiquitin. This will allow cleavage of the discrete polypeptides at translation similar to its successful application in diamondback moth (described in Chapter 4, section 4.3.10). In males the *Tctra* male exons will be included, preventing translation of the transcript (the male exons contain stop codons). In females this is spliced out, producing a functional transcript with part of the α -tubulin-1 exon and a chimeric exon of *Tctra* and Zsgreen that was included for ease of construction purposes (Figure 126). The ubiquitin should split these amino acids from the tTAV.

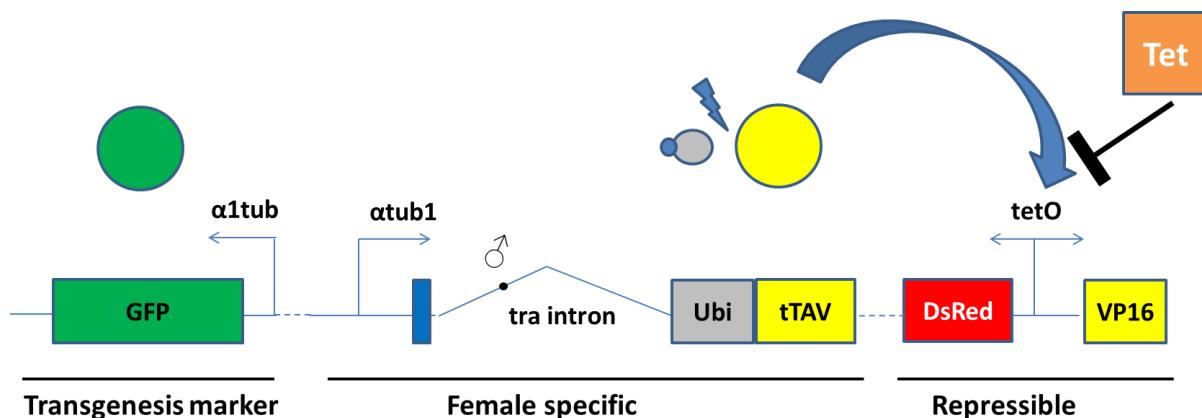


Figure 126. A general overview of the core components of a repressible female-specific expression system in *Tribolium* designated OX5202. The DNA is represented by lines and rectangles (for genes) with arrows for promoters. The α -tubulin-1 promoter also includes 5' UTR and an intron which is suspected to be important for transcription. The female-specificity is provided by the *Tribolium transformer* intron whereby stop codons (black circle) are not spliced from the male transcript. Females translate the chimeric ZsGreen and α tub exon2 (blue), ubiquitin (grey) and tTAV (yellow) which are cleaved at the ubiquitin (blue zap). The tTAV will then activate *tetO* in the absence of tetracycline (here tetracycline – Tet - blocks expression). Proteins are represented as circles and tetracycline as a square. Some components have been omitted for clarity.

5.4 CONCLUSIONS

Tribolium survived on a tetracycline-infused diet. Combined with previous literature reviews, which suggested that the antibiotic was successfully ingested and transported to the gonads of the insect (Wade and Stevens, 1985), the data lends support to the plausibility of developing a tetracycline-repressible paternal effect system.

The system was developed piecemeal, starting with confirmation of the endogenous *Tc α-tubulin-1* promoter as a suitable transformation marker when combined with DsRed2.

Following transformation we provided a more detailed description of the expression profile of this gene using multiple lines, verifying the work of Siebert *et al.*, (2008). This promoter provided sperm expression during spermatogenesis suggesting that the gene may be involved in microtubule formation in the sperm axoneme (the *Dm* orthologue *αTub84B* shares this function). This renders it unsuitable as a transformation marker in a paternal effect system if distinct sperm-marking is also desired.

Interestingly, *α-tubulin-1* was found in both the soma and the germline suggesting it is involved in microtubule formation in both cell types. This is similar to the Diptera but different from the Lepidoptera (Chapter 4). This suggests that the paternal-effect system in *Tribolium* could be based on the successful *β2-tubulin* system used with the Diptera.

A male skew in the sex-ratio was noted when collecting transgenic adults. This, combined with difficulty in screening at the pupal and adult stage, suggested that the marker needed to be optimised. The *Tcα-tubulin-1-DsRed2* combination performed poorly which may have been due to the inconspicuous nature of expression in pupal and adult females relative to the males. This implied that, as a transformation marker, we may be more likely to miss our female transformants (relative to males). This may explain the poor transformation of OX4799 (given the relative number of G₀ compared to other constructs). Following this conclusion, *Tcα-tubulin-*

1-ZsGreen was preferred and shown to be superior in marker reliability assays of later constructs.

The development of the female-lethal module also progressed after characterisation of *Tcdsx* and *tra*, the latter was selected as our gene of choice for the basis of a female-specific expression system. A transgene was successfully regulated using this method, producing correctly spliced transcripts specific to the female. In males, the stop codon-containing male exons were not spliced out and hence did not produce functional protein. However, despite apparently correct female transgene transcription and splicing, no functional DsRed2 was detectable in the females. This may have been due to interference of correct DsRed2-folding caused by the 5' end containing nine nucleotides of *Tctra* (as well as the 5' UTR and promoter region). This possible failure in folding was also seen in diamondback moth. The same remedy was attempted whereby ubiquitin was inserted between the problematic region and the DsRed2 to separate the amino acids chains post-translation, facilitating correct folding.

This project finished with a final construct design intended to combine the conditionality and sexing modules which should kill females off-tetracycline via a positive feedback loop of tTAV. Construction difficulties of OX5202 delayed synthesis by 6 months. This provided insufficient time to inject and analyse in time for inclusion in this thesis.

Sperm-specific expression was not attempted due to the pending success of the conditionality and female-specific modules of the paternal effect system yet to be completed.

5.5 FUTURE DIRECTION

Following successful transformation with OX5202 translation of tTAV will drive the *tetO* bi-directional enhancer, resulting in transcription of DsRed2 and VP16, respectively, both serving as reporter proteins. DsRed2 may be visible in the females as seen in OX4700 and the transcript of VP16 can be detected by RT-PCR. The inclusion of tetracycline in the diet should make this

part conditional, potentially repressing tTAV from activating *tetO*. This can be assessed by rearing transformants on and off tetracycline diets (Table 35).

Table 35. Expected outcome of rearing transgenics on and off tetracycline.

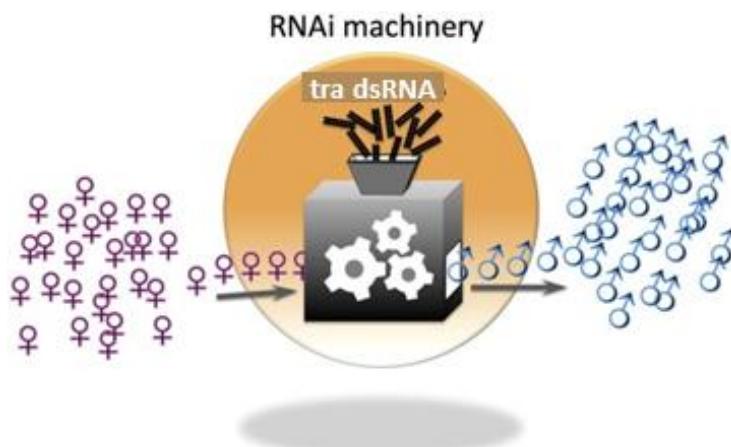
| On tetracycline | Off tetracycline |
|--|---|
| Green male and female, no DsRed2 or VP16 expression. | Green male and female, also female-specific DsRed2 and VP16 expression. |

Following successful development of *Tribolium* strains carrying a female-specific autocidal genetic system, the system should be transferred to other coleopteran pest species to assess its functionality. If successful, the system will provide potentially valuable alternative to conventional sterility-based methods while also eliminating females from the release population in pest coleopteran target species of interest.

As most agricultural pests cause economic damage mainly during the larval stage, research should continue in developing the paternal effect system in *Tribolium*, anticipating its transfer to other species that prove compatible with the female-specific system.

Currently both the paternal effect and the female-specific lethality would be made conditional on tetracycline presence. Combining these systems in one insect line may benefit from each system being conditional on orthogonal binary-systems. Research should focus on alternative binary systems for regulating the different genetic modules. As an alternative to using genetic binary systems and insect transgenesis, RNAi could be exploited to alter the sex ratio of insects produced during mass rearing. Work conducted to investigate this possibility is described in Chapter 6.

CHAPTER 6: INSECT TRA-GENDERESIS: SILENCING FEMALNESS AS PART OF A MASS-REARING PROGRAMME



6.1. INTRODUCTION

6.1.1 RNAI IN INSECTS

RNA interference (RNAi) has revolutionised insect-based research, enabling scientists to suppress a gene of interest thereby linking a phenotype to gene function (reviewed in Scott *et al.*, 2013). RNAi machinery is found in most eukaryotes. The experimental use of RNAi exploits this innate immunity against viruses and transposable elements. RNAi has potential for augmenting the genetic control of agricultural pests through incorporation into SIT programmes (Whyard *et al.*, 2015).

Downregulation of specific genes through RNAi has been widely used for insects (Price & Gatehouse, 2008; Whyard *et al.*, 2009; Terenius *et al.*, 2011; Shukla & Palli, 2012; Yu *et al.*, 2013; Singh *et al.*, 2013). The typical method of delivery of synthesised double-stranded RNA (dsRNA) is by microinjection into the insect. Injecting insects with dsRNA to achieve RNAi limits the scalability of the method if it is to be considered for large-scale biotechnological applications.

Interestingly, just like in *Caenorhabditis elegans* - in which bacteria expressing the dsRNA can be fed directly to generate phenotypic effects due to gene silencing - some insect genes have been silenced by direct feeding of dsRNA (Huvenne and Smagghe, 2010). The process includes the uptake of dsRNA by cells in the gut lumen of an insect inducing a systemic RNAi response. Common application approaches include transgenic plants expressing a RNA hairpin for the target gene (Mao *et al.*, 2007) or ingestion of artificial diet (Baum *et al.*, 2007; Asokan *et al.*, 2013) containing the dsRNA.

Mass-rearing insects during an SIT programme offers an opportunity to incorporate dsRNA into the artificial diet to regulate gene expression in insects to be released. This could be used to augment the released insects (McGraw and O'Neil, 2013) or to improve the economics of the programme by affecting mass-rearing efficiency.

Combining recent discoveries by Shukla & Palli (2012) with this strategy may result in a method of genetic sexing in the model coleopteran, *Tribolium castaneum* (Brown *et al.*, 2009), whereby insects could be reared on diet with dsRNA silencing *transformer* (*tra*). The sex determination system shares commonalities with other insects whereby *tra* and *transformer-2* (*tra-2*) regulate *doublesex* (*dsx*) splicing (Pease & Hahn, 2012). Research in this area will improve understanding of this common regulatory pathway.

Silencing *Tctra* in XX females has been shown to produce a pseudo-male phenotype. This strategy could be used as part of an SIT programme. It is attractive as it avoids the wasted labour, space and diet associated with feeding females that will never be released.

However, these pseudo-males are also much more likely to die during early embryonic development most likely because of mis-regulation of complex dosage-compensation genes (Shukla and Palli, 2012; Shukla and Palli, 2014). Perhaps it will be possible to develop an appropriate dosage to express *Tctra* above the threshold levels required to inhibit the activation of dosage compensation pathways but below the threshold levels required to execute the splicing of *Tcdsx* pre-mRNA into the female isoform. This could be used prior to the release generation to either kill or convert female offspring. Thus only sterile males and pseudo-males would be released.

6.1.2 HYPOTHESES

In Chapter 5 we aimed for a conditional female-specific-lethal genetic system in *Tribolium* using insect transgenesis. This Chapter provides an alternative genetic-sexing system; where we assess the viability of exposing *Tribolium* to dsRNA *Tctra* (ds *Tctra*), to attempt to induce maleness in females. These experiments were designed to test the following hypotheses:

- Injecting *Tribolium* with ds *Tctra* increases the risk of death in females compared to males.
- Injecting *Tribolium* with ds *Tctra* will induce the male phenotype in genetic females.
- Feeding *Tribolium* with ds *Tctra* will produce the same effects as injecting.
- We can assess the performance of the ds *Tctra* in regulating the sex ratio during mass-rearing by modelling using parameter estimates from this Chapter.

6.2 RNAI IN *TRIBOLIUM*

6.2.1 *TRIBOLIUM* AS A MODEL FOR RNAI

Some species are particularly amenable to RNAi, with *Tribolium* considered a model RNAi organism in the insects (Posnien *et al.*, 2009). Importantly parameters that affect the efficiency

of RNAi in *Tribolium* have been characterised, providing some guidance for experimental design (Miller *et al.*, 2012). The dsRNA has been delivered by both injection and feeding through diet to various life stages. Each has associated costs and benefits but given our eventual objective of creating all male cohorts for sterile insect release, feeding seems the more suitable for high through-put. If this method proves untenable then an alternative of transgenic expression of an RNA hairpin for *Tctra* could be developed, built on this research.

6.2.2 EXPERIMENTAL DESIGN CONSIDERATIONS FOR *TCTRA* RNAI

Prior to any RNAi experiment the success hinges on the design and production of a specific RNAi molecule for the target gene of interest, in this case dsRNA of *Tctra*. Transformer protein (Tra), along with Transformer-2 protein (Tra-2), regulates splicing of *dsx* pre-mRNA in females (Shukla & Palli, 2013). Tra is required for XX insects to follow the female sex determination pathway. Tra acts as a memory device for the female determination cascade through self-regulation of splicing. The original *tra* transcript translated in an embryo originates from the mother and is provided in the egg following fertilisation. In other words, injecting a female with dsRNA for *tra* induces RNAi and can result in all-male progeny, effectively silencing femaleness by destruction of the maternally inherited *tra* mRNA. Our experimental design is based on that of Shukla and Palli (2014).

In addition to injecting into female pupae and inspecting the offspring, it is also of interest to investigate how the silencing of *tra* affects development when induced early on. This is to be investigated by embryo microinjection to induce RNAi (Figure 127), an experiment not conducted by Shukla & Palli (2012).

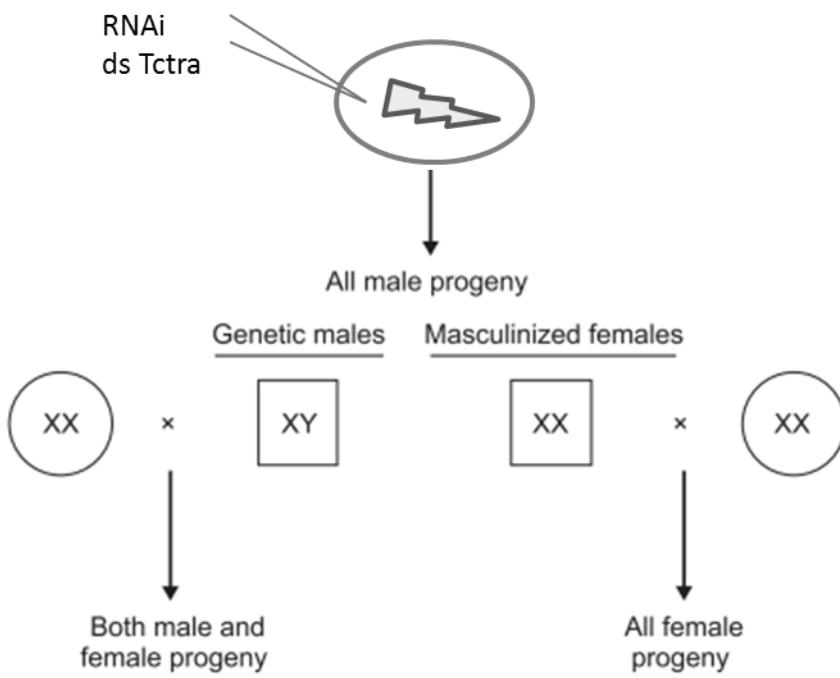


Figure 127. Schematic diagram summarising the experimental design and results predicted based on work by Shukla and Palli in their studies on sex determination in *Tribolium castaneum* elucidated with RNAi (2012; 2013; 2014). Squares represent the male phenotype and circles the female phenotype. Masculinised females lack the Y chromosome so are homogametic. When injecting larvae, Shukla & Palli noted the suppression of the development of pupal (finger-like papillae) and adult female (gain of male bristles on a small patch of the inside of the first pair of legs a third of the distance from the base) sex-specific structures.

Successful silencing may depend on the maternally inherited transcripts and whether they can be silenced before they are translated in the females. There is some evidence (qRT-PCR), described by Shukla & Palli (2012; 2013; 2014), that the transcript levels of the female-specific isoform (*Tctra-f*) are relatively high in the unfertilised eggs, falling after fertilisation to a third of the initial relative transcript abundance. A peak of *Tctra-f* is detected at 12-13 h after egg-laying and a peak of male-specific isoform (*Tctra-m*) is detected at 18-23 h after egg-laying. The effect of silencing *Tctra-f* in embryos, and the resultant pupal phenotype, was elucidated by microinjection.

6.2.3 EXPERIMENTAL TREATMENT AND CONTROL CONSIDERATIONS

We included *Tctra* dsRNA designed by Shukla and Palli (2012) as a positive control (treatment A), and also a shorter 200-bp sequence of our own design [treatment B (a dsRNA treatment of half the size was preferred to reduce mass-rearing synthesis costs)] to be more congruous with

our negative control (treatment C), which is dsRNA for the male-specific exons 2a and 2b of *tra* (size was limited by the length of the male-specific region). The male-specific dsRNA controls for both the administration of the dsRNA and the physiological impact of the RNAi cascade.

The dsRNA dosage should have been adjusted for transcript abundance and protein stability. The transcript relative abundance was known at the different life-stages, but protein levels and stability were not. The appropriate dose also depended on the mode of uptake. Fortunately, prior experimentation has led to the development for injection protocols for all life stages (Berghammer *et al.*, 2009; Posnien *et al.*, 2009), as well as success in diet-fed dsRNA-mediated RNAi in *Tribolium* larvae (Whyard *et al.*, 2009).

Given the three treatments, one may have expected an ordered sequence of efficacy. The longer dsRNA sequence (Shukla & Palli, 2012) was predicted to produce more siRNAs against the targeted mRNA transcript; potentially increasing the RNAi response. This was evaluated by the examination of sex-specific-morphs in the progeny following injection.

6.2.4 REPRODUCING AND BUILDING ON RESEARCH

Reproducing any RNAi work is important as it has been historically unreliable and variable between research groups (Yu *et al.*, 2013).

6.2.4.1 DSRNA DESIGN AND SYNTHESIS

The Shukla & Palli (2012) primers were used as the positive control (A; Figure 128) to create an appropriate cDNA template for RNA transcription. We designed our own primers to create a predicted positive (B) and a negative (C) for inducing silencing of *Tctra*.

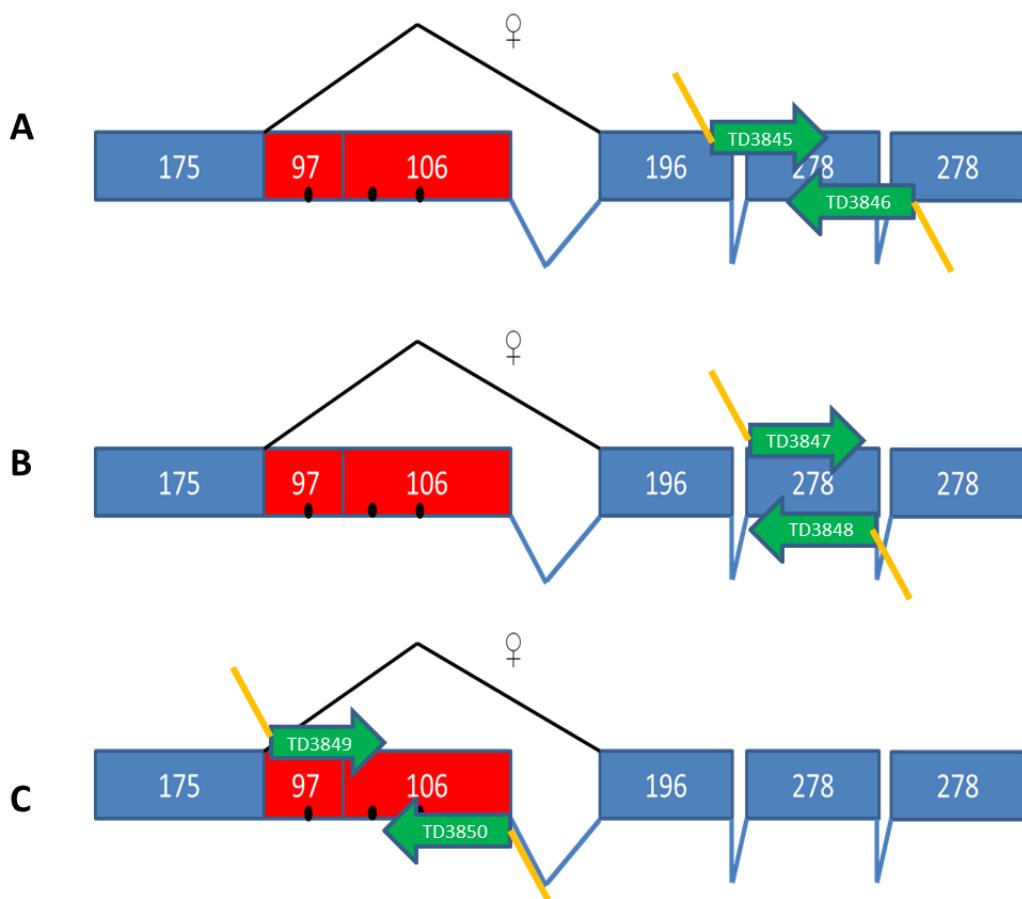


Figure 128. Schematic diagrams showing the three PCR targets of *Tctra*. Primer pairs with T7 promoter regions on the 5' end (orange line; 5'-GAATTGTATAACGACTCACTATAGG-3') were used to amplify, by high fidelity PCR, DNA that would be used as a template for *in vitro* transcription of dsRNA to be used for RNAi. The primers are shown as green arrows with the respective primer names. The exons of the *tra* gene are shown as boxes and introns as lines (shared introns – blue, female specific intron - black). The number of nucleotides per exon is shown. A) The positive control (primers TD3845 and TD3846) based on Shukla and Palli (2012) (401 bp), B) our novel design (TD3847 and TD3848) (201 bp), C) the negative control (TD3849 and TD3850), unlike Shukla and Palli our negative control (201 bp) is found within the same gene. The black circles represent stop codons in the male-specific exons, which results in truncated non-functional protein in males. The splice variation between the female (top) and male (bottom) isoforms are shown.

The primers with T7 promoter were synthesised and used in a PCR to amplify the target regions of the *Tctra* gene. The PCR products were purified in a column (Figure 129) and then RNA was transcribed *in vitro* and, following precipitation and heat treatment, ready for injection.

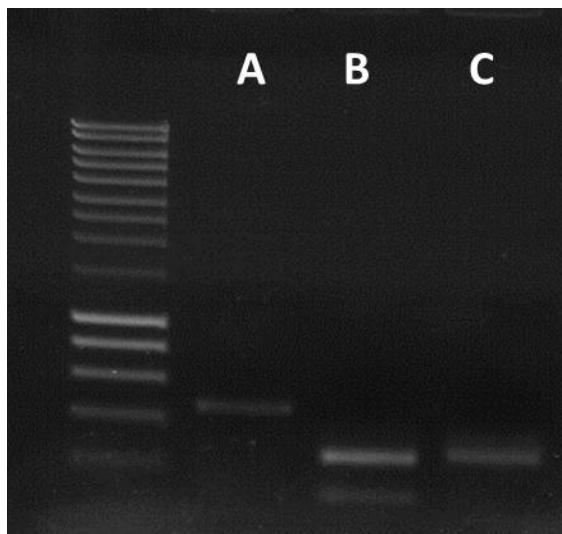


Figure 129. Synthesising dsRNA for RNAi of *Tctra*; from DNA to RNA. One μ l of each PCR product were resolved on a 1.2% agarose gel. Without the T7 promoter region amplified regions were expected to be A) 401 bp, B) 201 bp and C) 201 bp. The sizes were as expected and the concentration was adequate for RNA transcription. Primer dimer formation is visible in lane B. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder). Two μ l of each column purified PCR product were used with the MEGAshortscript T7 kit (Invitrogen, UK). Transcription reactions were run for 3 h at 37°C. The RNA was precipitated using Lithium Chloride and re-suspended in 40 μ l of nuclease-free water (using mMESSAGE mMACHINE Kit, Invitrogen, UK). The spectrophotometer was used to quantify the RNA concentration. The RNA concentration was deemed too low according to the protocol so was re-precipitated and eluted in half the volume. Preliminary injections revealed an injection volume per pupa of 0.5 μ l suggesting on average between 300-600 ng of dsRNA was injected. The RNA was heat-treated at 95°C for two min and then cooled from 95°C to 70°C over the course of 15 min.

6.2.4.2 TRANSFORMER RNAI IN *TRIBOLIUM CASTANEUM* PUPAE

6.2.4.2.1 INJECTIONS

Pupae were injected as per the protocol (Posnien *et al.*, 2009), with poor survival of the RNA treatments relative to a water control, putatively due to preparation toxicity (Table 36).

Table 36. DsRNA injections into *Tribolium* pupae to silence sex-specific transcripts of the *Tctra* gene. Survivors of dsRNA-injected individuals was relatively poor to the water only control. The three survivors of treatment B were sickly and still yet to eclose. The insects appeared to have survived the initial trauma with most dying pre-eclosion as very late pupa. This may suggest toxicity of a component specific to the dsRNA preparation or the effect of the dsRNA itself. A suitable control undergoing the same precipitation of the three treatments without the dsRNA was neglected thus leaving uncertainty as to the cause of the higher than expected mortality. It may have been due to inadequate removal of the Lithium Chloride during the RNA precipitation step. Chloride ions inhibit protein synthesis and DNA polymerase. Brackets provide survival as a percentage to one decimal place.

| Treatment | Number of injected pupae | Survivors (1 week) |
|----------------------|--------------------------|--------------------|
| A (positive control) | 29 | 0 (0) |
| B (our design) | 32 | 3* (9.0) |
| C (negative control) | 24 | 4 (16.6) |
| D (water only) | 34 | 25 (73.5) |

6.2.4.2.2 SECOND ROUND OF INJECTIONS

The dsRNA for each treatment was re-synthesised and precipitated this time airing the pellet at 37°C for 30 min to ensure thorough drying. An additional control of no dsRNA (essentially water subjected to the LiCl precipitation step) was included among the treatments to control for the effect of injecting with lithium chloride.

Table 37. DsRNA injections into *Tribolium* pupa to silence sex-specific transcripts of *Tctra* including a negative control of no dsRNA that controls for the precipitation of the dsRNA step and injection. Survival improved relative to the first batch of injections. The treatments were not significantly different from one another assessed with logistic regression. *The water-only treatment was not repeated with the results taken from the first experiment. The brackets show the percentage success to one decimal place.

| Treatment | Number of injected pupae | Survivors (1 week) |
|-----------------------------|--------------------------|--------------------|
| A (positive control) | 45 | 12 (27.7) |
| B (our design) | 45 | 10 (22.2) |
| C (negative control) | 38 | 8 (21.1) |
| No dsRNA (negative control) | 26 | 6 (23.1) |
| Water only* | 34 | 25 (73.5) |

The data was modelled using a logistic regression where success was an injection survivor at one week (the number of failures were those pupae that did not survive injection). The no-dsRNA-injected embryos were considered the control treatment and were used as the baseline or reference category for the regression.

Taking the exponential of the coefficients of the model we can say that all treatments that involved the lithium chloride precipitation step were not significantly different in terms of their odds ratio of success from one another. This is indicated by the 95% CI containing one for treatment A (mean odds of 1.21 and 95% CI of 0.40-3.95), B (mean odds of 0.95 and 95% CI of 0.31-3.16) and C (mean odds of 0.89 and 95% CI of 0.27-3.07). This can be interpreted as treatment A averaging 1.21 successes for every one success of the No dsRNA treatment. This suggests that the dsRNA itself is not having a significant effect on pupal survival.

In contrast the water-only control was significantly different from the no-dsRNA control (mean odds of 9.26 and 95% CI of 2.97-32.75, $p < 0.001$), with much improved survival. This indicates toxicity associated with the precipitation methods, supported by the conclusions of the first round of pupae injections.

Female injection survivors were crossed with one wild-type male on diet. The flour was collected after 20 days, and the offspring were allowed to develop to pupation (some insects had eclosed), when the sex of the offspring was determined (Table 38).

Table 38. DsRNA injections into *Tribolium* pupa to silence sex-specific transcripts of the *Tctra* gene produced surviving females that were crossed with wild-type male. The sexes of the offspring were recorded. The averages for the fertile productive crosses are shown to the nearest integer. Unfortunately the high mortality of the RNAi treated females continued after eclosion with death early in adulthood. The crosses were inspected and indeed it was typically the female that had perished. In one cross with A, and two with C, the male had died. These females were set up in a new cross but the data was excluded from this analysis as the time since injection would be different. The SE is not shown as the variation is described graphically in (Figure 130).

| Treatment | No. productive crosses | No. crosses not producing progeny | Males | Females |
|--------------------------------|------------------------------|---|-------|---------|
| A (positive control) | 1 | 11 | 45 | 0 |
| B (our design) | 0 | 10 | NA | NA |
| C (negative control) | 2 | 6 | 20 | 24 |
| No dsRNA (negative control) | 9 | 2 | 23 | 27 |

These data can be displayed as a scatterplot where each replicate is plotted (Figure 130).

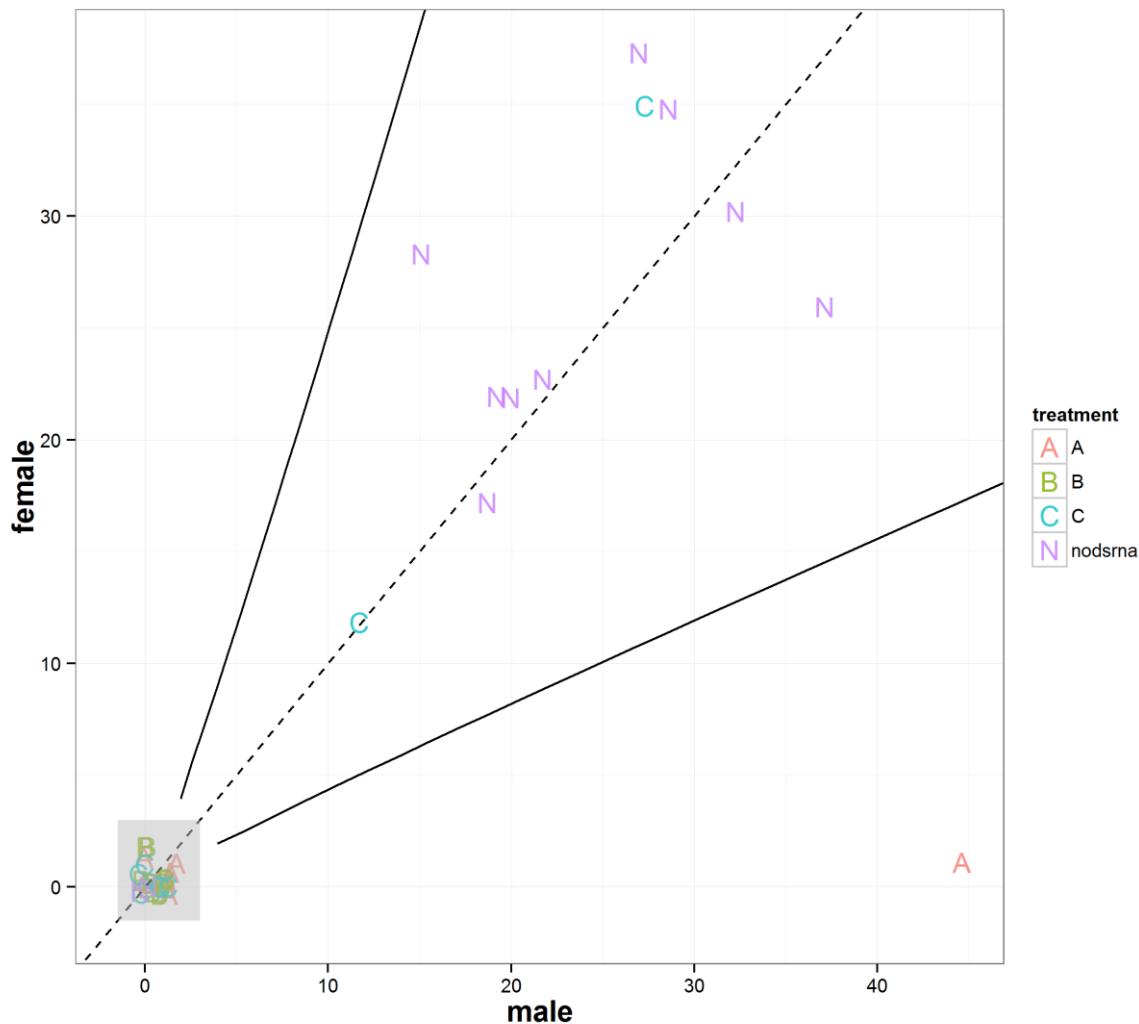


Figure 130. A jittered-scatterplot of the male and female progeny produced by each cross coloured and shaped by treatment. Each replicate provided a male (x) and female (y) count as the response variable. A caveat, each datum can be out from reality by up to one unit due to the jitter. Comparing this to the expected 1:1 sex ratio suggests a pattern of distinct outcomes to the treatments. Three types of outcome: the females are killed post-eclosion due to the *Tctra* dsRNA (or dsRNA immune response), or there was no effect (the negative controls that survived to produce progeny) or all male offspring were produced by silencing the maternally inherited *Tctra* transcripts. Most of the dsRNA-injected crosses failed to produce progeny due to female death or more rarely infertility. Male death in the cross prevented some of the crosses from producing offspring but this was rare. These groups are described by the grey square region. Of the negative control treatments, C was less likely to survive but when the female produced progeny they were not dissimilar from 1:1 and the other negative control of no-dsRNA-injected. The final distinctive group is that of the sole female to produce progeny from treatment A which produced only male offspring. The treatments had no effect on the sex ratio compared to the null model (1:1 dashed line) except for one replicate of treatment A. The 95% binomial confidence upper and intervals of the null model are shown as solid lines (the Wilson method was used).

To assess the patterning of the sex ratios and whether my intuition that there were three distinct groups or clusters in the data following RNAi treatment, K-means clustering was used to objectively (it is reproducible given the random seed) partition the data into K distinct, non-overlapping clusters. Essentially we wish to identify sub-groups within the data, with the three

groups representing the non-fertile crosses, crosses yielding progeny with a 1:1 sex ratio and crosses yielding all males. Prior to this analysis the data were corrected by removing one from each sex from each cross to a minimum of zero (this accounts for the original parent still being present in the cross on counting). The K-means clustering confirms intuition separating the data almost perfectly into the three putative groups.

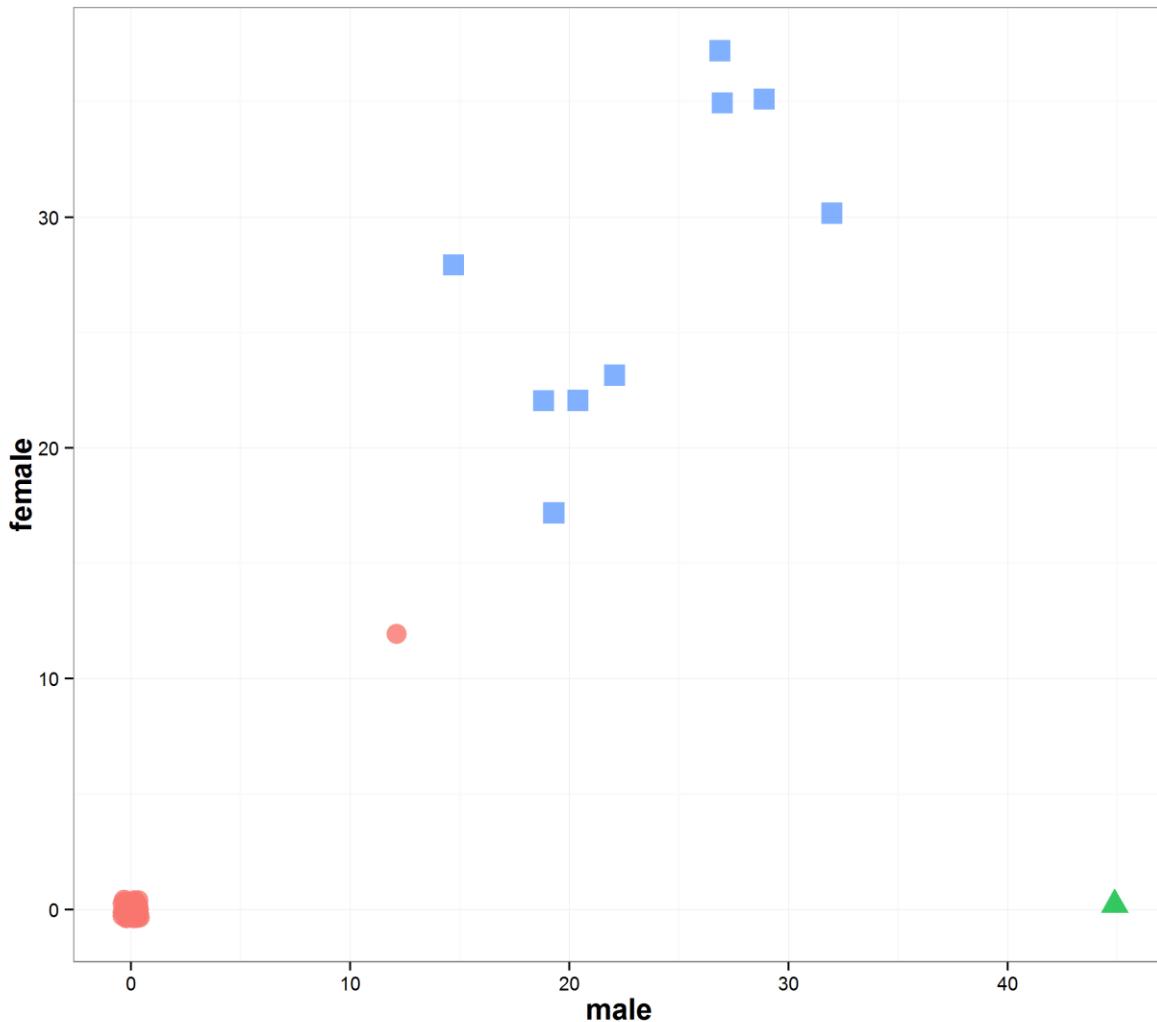


Figure 131. The clustering data set involving hypothesised three groups. Each group is shown using a different coloured symbol. The three groups show good separation except for the treatment C datum at coordinate (12, 12) which is misclassified. The insects for this cross developed more slowly with 12 larvae present when sampled for sex ratio. This supports the conclusion that the negative controls either produce non-viable crosses at a relatively low rate or produce 50:50 sex ratios (blue squares). The positive controls either give rise to dead females or non-viable crosses most of the time (red circles) or rarely give rise to crosses that produce just males (green triangle).

Shukla and Palli (2012) had shown that *Tctra* knockdown in larvae affects mortality both immediately following injection and after eclosion compared to the control. However, the

toxicity we experienced seems above that experienced by their group. We did achieve one successful cross of the positive control (treatment A), where success was all-male progeny.

For the negative control treatments, male-exon-targeting and no dsRNA, all the phenotypic males were confirmed as true males by genotype (the male offspring of one randomly selected cross for each treatment were examined with 20 males for each confirmed as true males). The females were confirmed as true females with 10 insects genotyped per control treatment (Figure 132).

| | | Phenotype | | |
|----------|---|-----------|----|----------|
| | | ♂ | ♀ | |
| Genotype | ♂ | 39 | 0 | 39 |
| | ♀ | 3 | 0 | 3 |
| | | 42 | 0 | 42 |
| A | | ♂ | ♀ | |
| | | NA | NA | NA |
| | | NA | NA | NA |
| | | NA | NA | NA |
| B | | ♂ | ♀ | |
| | | 20 | NA | 20 |
| | | 0 | NA | 0 |
| | | 20 | NA | 20 |
| C | | ♂ | ♀ | No dsRNA |
| | | 20 | 0 | 20 |
| | | 0 | 10 | 10 |
| | | 20 | 10 | 30 |

Figure 132. The phenotypic sex of the pupal offspring compared to the genotypic sex by RNAi treatment. Treatment B produced no offspring to sample. The negative controls showed no evidence of pseudomaleness generation. The offspring from the cross that produced just males (the only productive treatment A derived cross) were dissected at adulthood and testes examined for proper formation and scored accordingly. Testes were found in all but one insect this may have been due to human error (this was genotyped as male later). The males were then genotyped using the multiplex-sexing PCR method to determine the presence of the Y chromosome. Due to the low number of pseudomales no relationship between pseudomaleness and testes quality was detected compared to true males (data and analysis not shown). Despite the number of offspring in the treatment A cross being similar to the other treatments it had a heavy XY skew. We expected a 1:1 distribution of XY and XX pseudomales based on Table 38.

The genotyping proved surprising, revealing a low occurrence of pseudomales. Again, perhaps this can be explained by dosage compensation interference caused by the double X chromosome and lack of *tra* regulation. The increased gene expression could result in death of most of the pseudomales before the pupal stage.

Given the pseudomale rate of 3/42 (7%) we lacked the number of males in our negative controls to detect a pseudomale. For example, the probability of seeing no pseudomales in n males would be:

$(1-(3/42))^n$ or as a function of n, where n is the number of males genotype (Figure 133).

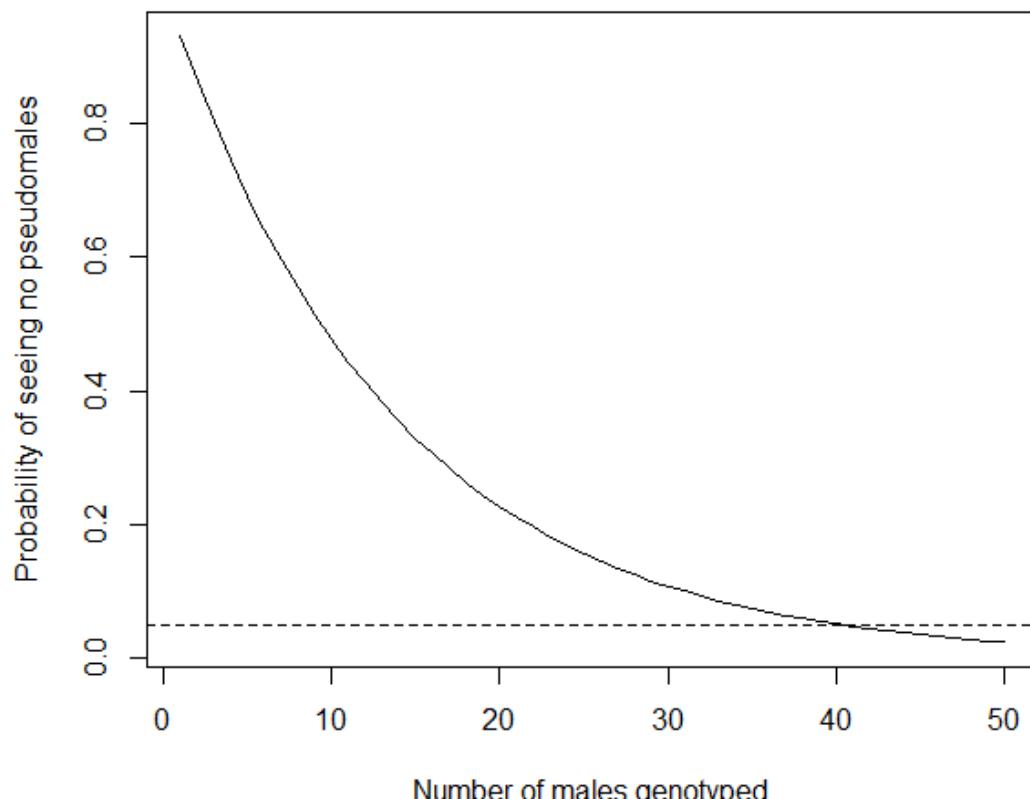


Figure 133. Power analysis revealing our lack of power in rejecting that no pseudomales are produced by the negative controls. Hence more replicates would be needed to make this conclusion at the 95% confidence level. This assumes a pseudomale generation rate equivalent to Treatment A (7%).

Following this interpretation we genotyped up to 40 males for the negative controls and found no pseudomales. This confirms with more confidence that only treatment A when injected into female pupae produced pseudomales in the next generation.

6.2.4.3 *TRIBOLIUM CASTANEUM* TRANSFORMER RNAI IN EMBRYOS

6.2.4.3.1 INJECTIONS

The principle of dsRNA insertion into the embryo was identical to that used for germline transformation microinjection (post injection temperature of 26°C). Again the survival was low (Table 39).

Table 39. DsRNA injections into *Tribolium* embryos to silence sex-specific transcripts of the *Tctra* gene. The larvae were reared to pupae and sexed by phenotype. The DNA was then extracted and genotyped for sex to confirm genetic sex and compare against phenotypic sex. The survival was very poor even for the water only control. The numbers in brackets represent the percentage survival between stages.

| Treatment: | A (+ve control) | B (our design) | C (-ve control) | D (water only) | E (no inject) |
|------------|-----------------|----------------|-----------------|----------------|---------------|
| Injections | 188 | 611 | 366 | 228 | 213 |
| Larvae | 6 (3.2) | 2 (0.3) | 3 (0.8) | 7 (0.3) | 95 (44.6) |
| Pupa | 6 | 0 | 1 | 5 | NA |

Survival falls well short of that expected by previous microinjections of *Tribolium* embryos with *piggyBac*. This may be due to the recommended approach of injecting as much volume as possible into the embryo to ensure silencing. Unfortunately a dead insect is not particularly informative.

No evidence of RNAi was found; all genetic sexes (assessed by multiplex PCR) matched the phenotype (Figure 134).

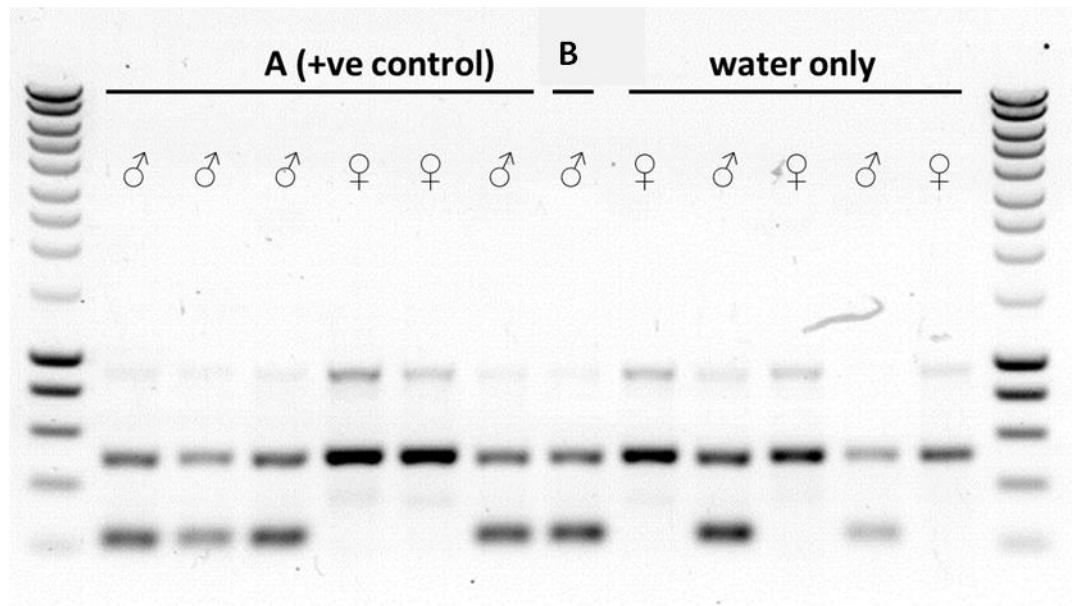


Figure 134. No evidence of *tra* silencing. The genotype of the *Tribolium* was compared to the phenotype by microscopy at the pupal stage. The PCR genotype is assumed to be the gold standard sex test using the multiplex PCR system developed by Lagisz *et al.* (2010). All genotypes matched the sex assigned by visual check of the phenotype. A ~500 bp band was detected in males and females, a ~230 bp band was found exclusively in PCRs with male DNA as expected. Treatment A was dsRNA that had been shown by Shukla and Palli to successfully silence *tra* in females. Treatment B was our design in attempt to achieve silencing with half the length dsRNA. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

The results could be explained by physical trauma killing any embryos injected with sufficient dsRNA to cause RNAi. Perhaps the only embryos to survive were those that were inadvertently not injected or injected with a small volume. The mortality for injection was unusually high compared to injections with *piggyBac* (Chapters 2 and 5), the difference may be due to the lack of viscosity resulting in too much volume being injected or the injection mix preventing wound closing of the embryo.

6.2.4.3.3 SECOND ROUND OF INJECTIONS

Repeating the dsRNA synthesis, precipitation and injection resulted in improved survival relative to the first round of injections (Table 40). Injection survivors were sexed at pupae and again at adulthood (sexing methods provided identical phenotypic outcomes), and a leg was removed for genotyping by multiplex PCR. An error rate of 1/48 for the genotyping method was

estimated by Lagisz *et al.* (2010). Appropriate blind control were therefore important to allow comparison.

Table 40. DsRNA injections into *Tribolium* embryos to silence sex-specific transcripts of the *Tctra* gene. The numbers in brackets represent the percentage survival between stages. The no-injection data was from the previous round of injections. The data could be interpreted as the discrepancy of females compared to the males in a treatment providing a probability of female dying due to toxicity. For treatment B we produced 66 males but only 26 females, a discrepancy of 40. This suggests that the mortality rate is approximately 40/66 (0.61).

| Treatment: | A (positive control) | B (our design) | C (negative control) | No dsRNA | No injection* |
|-------------------|----------------------|----------------|----------------------|-----------|---------------|
| Injections | 484 | 610 | 399 | 334 | 213 |
| Larvae | 38 (7.9) | 98 (16.1) | 154 (39.0) | 55 (16.5) | 95 (44.6) |
| Pupae | 32 (84.2) | 93 (94.9) | 140 (90.9) | 50 (90.9) | NA |
| Phenotype | 19 | 66 | 67 | 21 | NA |
| Male | | | | | |
| Phenotype | 11 | 26 | 73 | 28 | NA |
| Female | | | | | |

The data was visualised using a Sankey diagram constructed using the riverplot package in R (January Weiner, 2014). This reveals the relative losses at each stage and informs as to the sampling effort by treatment (Figure 135).

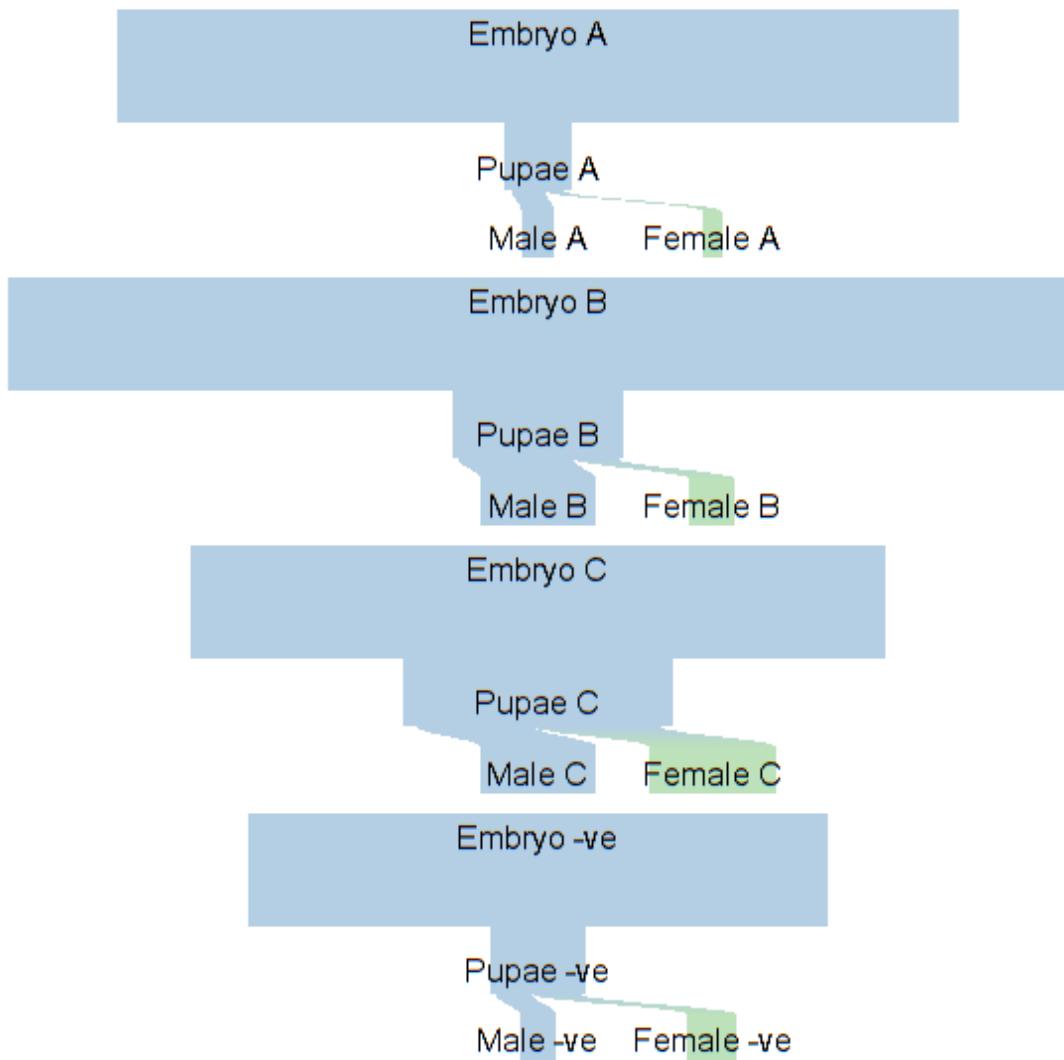


Figure 135. The Sankey plot shows the flow of the insects through life stages from injected embryo to sexed pupae. The width of the bar is proportional to the number of insects with the maximum-width-bar of embryos injected with treatment B at 610. Our design B was more heavily sampled as it was novel and not experimentally tested previously. The sampling effort was based on an expected survival to the pupal stage of 10% inferred from our *piggyBac* microinjection data, Chapter 2. Each treatment is shown, designated by the label suffix. Following the logistic regression this visualisation can aid interpretation. In this instance ‘-ve’ corresponds to the No-dsRNA treatment to avoid label clash.

A logistic regression model was fitted to the survival data from injected embryo to pupae. The no-injection data was ignored. The no-dsRNA-injected embryos were considered the control treatment and were used as the baseline or reference category for the regression. A success was an injected embryo surviving to the pupal stage and a failure was death as an embryo or larvae following injection.

6.2.4.3.3.1 SURVIVAL

Taking the exponential of the coefficients of the model (transforming them from log odds to odds) we can say that treatment A more than halves the odds of survival relative to the control (mean odds of 0.40 and 95% CI of 0.25-0.64), treatment B was equivalent to the no-dsRNA control (mean odds of 1.02 and 95% CI of 0.71-1.49) and treatment C almost trebled the chance of survival (mean odds of 3.07 and 95% CI of 2.15-4.55). If the 95% confidence interval includes the value '1' then the treatment in question is not significantly different from the control of no dsRNA injection.

For the convenience of the reader, the odds of success for each treatment were also transformed into percentages of successful injection survival to pupal stage (Table 41).

Table 41. The percentage probability of an embryo injection survival to pupae by treatment. Predicted from the logistic regression: survival~treatment.

| Treatment -> | A (+ve control) | B (our design) | C (-ve control) | No dsRNA |
|--------------------------|------------------------|-----------------------|------------------------|-----------------|
| Probability -> | 6.6% | 15.2% | 35.1% | 15.0% |

RNAi has often been associated with reduced survival due to an immune response and associated physiological costs. However, the treatments were designed to affect the femaleness of the injected embryos. The phenotype and genotype of survivors by treatment were compared (Figure 136).

6.2.4.3.3.2 PHENOTYPE VERSUS GENOTYPE FOLLOWING RNAI TREATMENT

| | | Phenotype | | |
|----------|---|-----------|----|----|
| | | ♂ | ♀ | |
| Genotype | ♂ | 18 | 0 | 18 |
| | ♀ | 0 | 9 | 9 |
| | | 18 | 9 | 27 |
| | | ♂ | ♀ | |
| C | ♂ | 29 | 0 | 29 |
| | ♀ | 0 | 32 | 32 |
| | | 29 | 32 | 61 |
| | | ♂ | ♀ | |
| No dsRNA | ♂ | 46 | 0 | 46 |
| | ♀ | 1 | 14 | 15 |
| | | 47 | 14 | 61 |

Figure 136. The phenotypic sex of the injection survivor compared to the genotypic sex by RNAi treatment. Fortunately one pseudomale was produced, vindicating the design of dsRNA treatment B. The positive control, treatment A, failed to produce any pseudomales, the associated probabilities of this are discussed in the text. Examination of the sex of pupae with a chi-square test revealed a non-significant departure from 1:1 sex ratio in all the treatments except B, however this approach may be inappropriate given the small numbers involved.

To be conservative the data was modelled using a logistic regression where success was a male phenotype. Of interest was whether any treatments affected the odds of maleness in injection survivors at the pupal stage.

Taking the exponential of the coefficients of the model we can say that treatment A doubles the odds of maleness relative to the control (mean odds of 2.30 and 95% CI of 0.92-6.00), treatment B was more than three times more likely to produce male pupae compared to the no dsRNA control (mean odds of 3.38 and 95% CI of 1.65-7.08) and treatment C (the dsRNA designed to elicit no response) was equivalent (mean odds of 1.22 and 95% CI of 0.64-2.38). Inspection of the 95% confidence intervals reveals the mean odds of a treatment to be misleading, although one could argue that treatment A is close to significance ($p = 0.08$), if it were not for the smaller number of replicates.

Treatment B (there was also weak evidence for A) was the only to differ from the control, evidenced by the 95% CI of the odds not containing one. This suggests it distorts the sex ratio, putatively by converting females to males or by selectively being more lethal to females. These results suggest that those treatments designed to invoke an RNAi response (silencing of *tra*) increase the odds of a male phenotype following injection into embryos (or increase female death disproportionately).

Table 42. The percentage probability of an injection survivor pupae being male by treatment predicted from the logistic regression: survival~treatment.

| Treatment -> | A (+ve control) | B (our design) | C (-ve control) | No dsRNA (-ve control) |
|--------------------------|------------------------|-----------------------|------------------------|-------------------------------|
| Probability -> | 63.0% | 71.7% | 47.9% | 42.9% |

The genotypes were inspected by leg PCR to determine whether treatment A and B were causing maleness in females. The other treatments were inspected to control for potential genotype test error rate.

6.2.4.3.3.3 LEG PCR MORTALITY

Although this method isolated DNA suitable for PCR from a small body part (Methods 8.2.23), leg amputation is an invasive procedure that likely affects fitness. Since the goal of non-lethal genotyping is to select animals with which to establish a cross, beetles must not only survive genotyping but also remain robust and fertile. The survival of these beetles after 2 weeks was compared to control non-amputated beetles reared under similar conditions after a dummy surgery.

Almost 60% of the pupae were genotyped and compared to their phenotypic sex (185/315). Of those insects, 74 were not genotyped (but were subjected to dummy surgery), only four of these died (5.5%) providing a control against leg removal. The insects that were genotyped by leg PCR suffered the trauma associated with leg removal with 24 of 236 dying (10.1%). Logistic

regression showed the difference in mortality between beetles with their leg removed against the control to be non-significant ($p = 0.22$).

6.2.4.3.3.4 INJECTION SURVIVORS CROSSED WITH WILD-TYPE FEMALES

All but one pseudomale from treatment B showed agreement between phenotype and genotype. To ascertain whether the pseudomale status was real (XX male), the pseudomale was crossed with a wild-type female. Other crosses were set up from the different RNAi treatments for comparison (randomly selected males were each crossed with a wild-type female). Mating pairs were allowed to mate for 20 days post-eclosion and resulting larvae were separated and allowed to mature to pupation for sexing (Table 43).

Table 43. Following dsRNA microinjection of a given treatment those embryos that developed to adulthood and had a male phenotype were crossed to virgin wild-type females following leg PCR. The one pseudomale produced from the experiment is the main point of interest with no male offspring produced suggesting that a factor responsible for male sex determination is present on the Y chromosome of *Tribolium castaneum*. The status also provides the G₀ survivor unique ID number for reference purposes. Individual crosses with *Tribolium* have high mortality due to beetles' inability to self-right.

| RNAi treatment | Status | Male offspring | Female offspring | Total |
|-----------------------|---------------------------------|-----------------------|-------------------------|--------------|
| A | male (G ₀ 55) | 28 | 7 | 35 |
| A | male (G ₀ 39) | 14 | 13 | 27 |
| B | pseudomale (G ₀ 182) | 0 | 32 | 32 |
| B | male (G ₀ 106) | 19 | 20 | 39 |
| B | male (G ₀ 147) | 28 | 18 | 46 |
| C | male (G ₀ 203) | 0 | 0 | 0 |
| No dsRNA | male (G ₀ 20) | 0 | 0 | 0 |
| Not injected | male (NA) | - | - | - |

Comparing sex ratios reveal the pseudomale to produce all female offspring; this confirms the genotype status as female despite the male phenotype. The pseudomale lacked a Y chromosome therefore only XX gametes were possible hence all-female offspring. This compares to crosses with true males that produce both male and female offspring. As the production of a pseudomale is rare we only had one replicate on which to test this.

Shukla and Palli (2014) described how all males were crossed with virgin females with 2.5% of crosses producing all female progeny (5/190 crosses). This is comparable to our data where 2.1% of treatment B microinjected survivor males were pseudomales (1/47).

For treatment A no pseudomales were produced in 18 males. Given the zero-numerator (Ludbrook & Lew, 2009) and the number of trials, the Bayes binomial 95% credible interval was 0-10% (Doraj-Raj, 2014), this suggests that we did not have the sample power to rule out that treatment A does not produce pseudomales ($10.0\% > 2.5\%$). This lack of power was due to higher than expected mortality following microinjection with treatment A. Furthermore, our data suggest that the masculinised genetic females (XX) are fully fertile, as the number of progeny produced when mated with virgin females is equivalent to the number of progeny produced in true male control crosses.

For completion, each pseudomale's female offspring was crossed with a wild type male to determine whether the sex ratio normalised (XX female and XY male parents). Both male and female offspring were detected in the next generation.

6.2.4.4 FEEDING DSRNA TO *TRIBOLIUM CASTANEUM*

There is experimental evidence that insect RNAi at certain developmental stages will affect other stages (Tomoyasu & Denell, 2004). Results from the literature (Shukla & Palli, 2014), combined with our findings, suggest that systemic RNAi of *transformer* is possible in *Tribolium*, albeit unlike the systemic RNAi seen in *Caenorhabditis elegans* (Tomoyasu *et al.*, 2008). Can this

be exploited in a mass rearing setting? Is it possible to silence the gene in adults via feeding of dsRNA *Tctra* with the aim of all-male production (Figure 137)?



Figure 137. A *Tctra* feeding bioassay to assess silencing in adult females prior to egg lay. Each treatment was randomly assigned to a cell with five replicates per treatment. The experiment was replicated in an additional 25-well plate. After this photo was taken coarse chaff was added to help the beetles self-right and avoid starvation.

Wells containing a shallow layer of diet (5 mm deep, $2 \times 2 \text{ cm}^2$) were overlaid with control reagents or dsRNA for *Tctra* in 10 μl volumes (allowed to dry for 30 min) at 0.3-0.5 $\mu\text{g}/\mu\text{l}$, and the females exposed (approximately 5000 ng per $4 \text{ cm}^2 \sim 1250 \text{ ng/cm}^2$). Seven days later (RNAi was shown to reduce target mRNA within a day of diet exposure in *Diabrotica virgifera virgifera*; Baum *et al.*, 2007 but our freshly eclosed females take a week to mature) a virgin male was added. After 1 week both insects were removed and additional diet added up to two thirds capacity of the well.

Two experiments were set up to vary the depth of diet added during the initial female exposure to the dsRNA. The above experiment was replicated with one change. The initial quantity of flour to which the dsRNA was added to was reduced so it barely covered the surface of the well-bottom (approximately 1-2 mm deep). This was an attempt to increase the ratio of treated flour to non-treated flour in a well, improving the probability of a female feeding on treated flour. The well was topped up with flour with the addition of the male.

After both adults were removed the eggs were allowed to develop under normal rearing conditions. Larvae produced from the crosses were allowed to mature and were sexed as adults by phenotype. The treatments had no effect on the sex ratio of the dsRNA-fed offspring. The data were combined below for convenience (Figure 138).

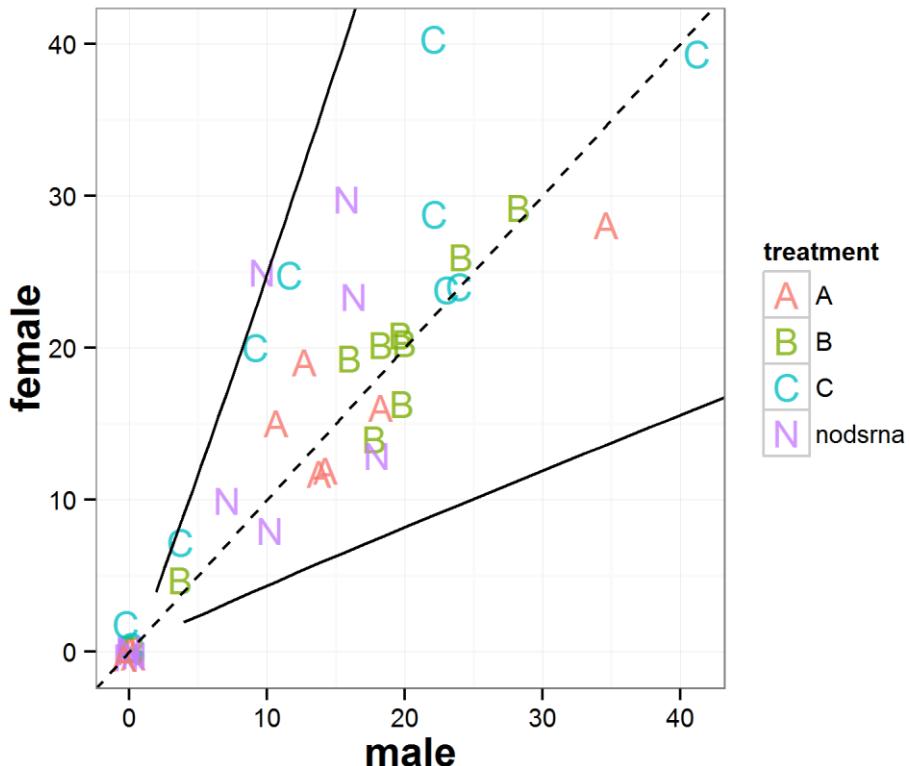


Figure 138. A jittered scatterplot of the sex ratio of offspring produced by adult female beetles treated with dsRNA in their diet. The treatments had no effect on the sex ratio compared to the null model (1:1 dashed line). The 95% binomial confidence upper and intervals of the null model are shown as solid lines (the Wilson method was used). There are ten replicates per treatment with some zero values. The reader should be wary of the random jitter added to the points and the risk of type 1 statistical error. Some of the crosses produced either very few or no offspring in the allotted time. Most of the replicates do not vary significantly from the null model based on an individual replicate inspection (the null model has not been corrected for multiple tests so we would expect some significant results due to chance). However, the average effect on treated females' offspring may elucidate a significant deviation from the no dsRNA control.

The data was modelled using a generalised logistic regression with binomial error distribution where success is the production of a male, failure a female. The GLM revealed that pooling the data suggested a significant difference of the no dsRNA control from a 1:1 male female sex ratio. The 95% CI of the odds of the treatment producing a male were 0.52 to 0.94, just shy of capturing one ($p = 0.02$). The other treatments were not significantly different from the no dsRNA control (A where $p = 0.06$, B where $p = 0.06$ and C where $p = 0.75$). Interestingly the two negative controls (no dsRNA and C) were closer in the estimate of the odds of producing a male and the two positive controls (A & B) were just shy of significance.

Also of note is the female skew of the offspring in some replicates of the negative controls: a chance event, perhaps, or it could be related to the phenomenon of sex adjustment of offspring seen in some insects. It is not known whether females of *Tribolium castaneum* are able to control the sex ratio of their offspring. Previous studies have found the average sex ratio to be close to unity although sometimes either significantly female or male biased (Sokoloff, 1974). However, there is some support for differential sex allocation in the literature albeit a weak effect, if real (Edvardsson & Arnvqist, 2005).

6.2.5 MODELLING EFFICACY

Given the data we have empirical estimates for the probability of female survival and pseudomale generation when applying the RNAi treatments to the insects. The feeding bioassay had no effect but let us assume that a similar efficacy can be achieved as with injection. The embryo injection data gave the probability of an injection surviving being male as 47 / 61 (0.77). We can estimate the probability of a female surviving the *Tctra* toxicity as 0.39 (1-0.61) (Figure 135). The probability of these females developing as pseudomales was 1/47 (0.021).

Assuming we are rearing 1000 insects, half of these will be female. Of these 500 females we expected only 0.39 to survive due to *Tctra* toxicity. Of these 195, on average four would be pseudomales. We can simulate the expected number of pseudomales by co-opting the “Goldilocks” model of Chapter 2 (Figure 139).

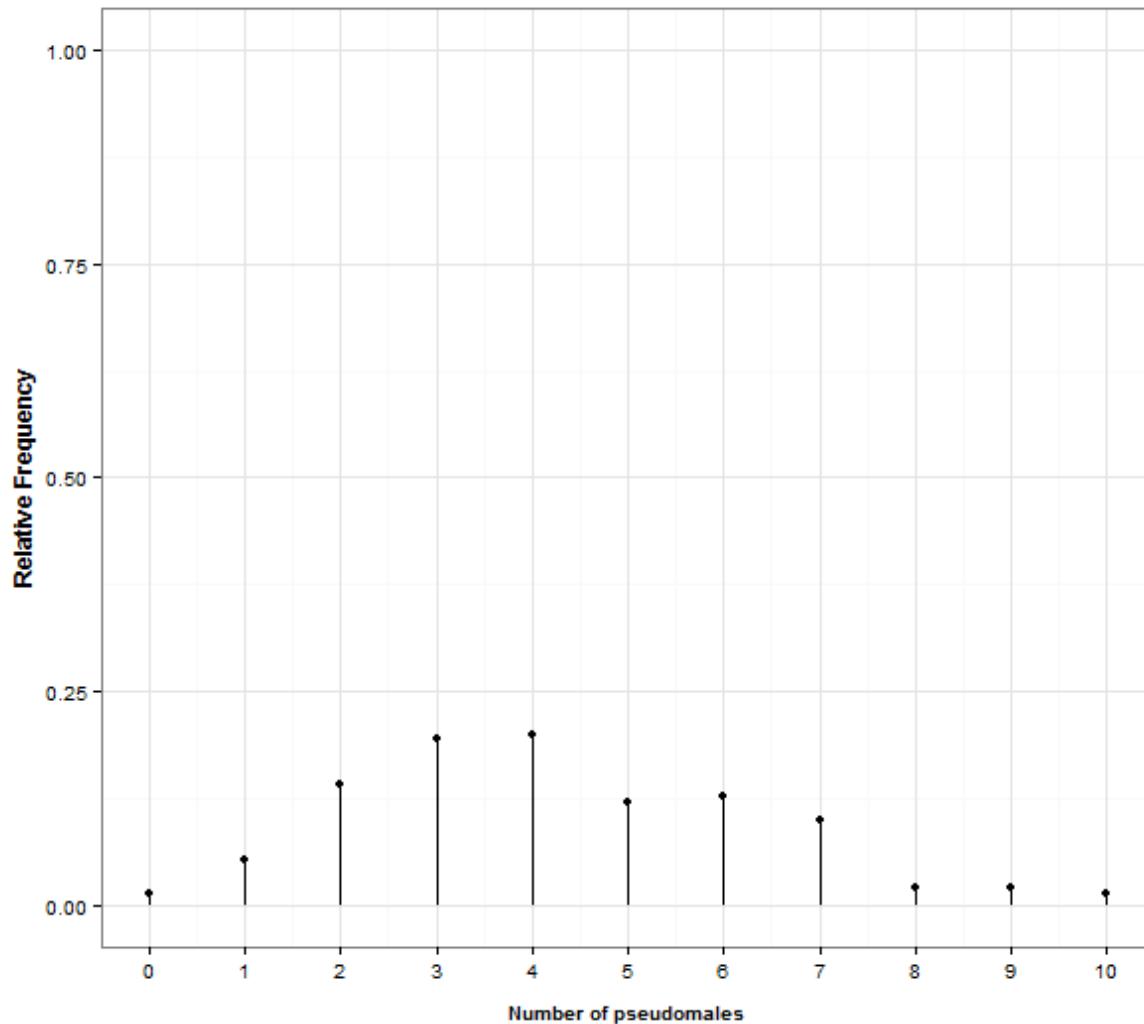


Figure 139. Given 500 females the number of pseudomales produced is simulated. The mode is four, with 95% CI approximately 1-8 pseudomales. This used the Goldilocks model with adjusted settings, inj=500, s=0.39 and g0=0.021. The rate at which pseudomales are generated is underwhelming with fewer than ten, given 500 females treated in this hypothetical example.

6.3 CONCLUSIONS

As more insects' genomes are sequenced, RNAi will continue to be a useful tool to ascribe functions to newly identified genes. The application of RNAi for pest insect control is set to expand with potential for release-insect augmentation or mass-rearing process optimisation using RNAi as part of an SIT programme. This study initiates the exploration of this area by using RNAi in *Tribolium* to adjust sex ratios and rearing efficiency by converting females into pseudomales.

DNA targeting the gene encoding *transformer* has been shown to affect the sex determination pathway in *Tribolium* in a previous study (Shukla & Palli, 2014). Our results verified their findings and improved the breadth of experiments employed by assessing RNAi of the *tra*-f isoform in embryos, pupae and using feeding bioassays in adults. We demonstrated similar efficiencies as described by Shukla & Palli (2014).

Despite success - silencing of *Tctra* - of our own novel dsRNA design and that used by Shukla & Palli in silencing *Tctra* we found that female mortality post-injection was relatively high compared to the males and the rate of pseudomaleness was very low (~2%). We support the conclusions of Shukla & Palli before us: most XX embryos exposed to *Tctra* RNAi are likely to die during early development as the second X chromosome is not inactivated. This affects or causes the misregulation of complex dosage-compensation genes.

Interestingly our microinjection of embryos produced similar effects as when injecting into later life stages. This suggests that translation of the maternally inherited *Tctra* transcripts occurs early after fertilisation as injected embryos are less than 6 h old. Enough transcript is translated to provide some Tra protein facilitating X-inactivation and dosage compensation in females.

Our data suggest that the pseudomales (XX) are fully fertile, as the number of progeny produced by the virgin females mated with these males is almost equal to the number of progeny in crosses between control beetles. This supports their application in an SIT programme as the pseudomales have the potential to compete for fertilisations. Furthermore, this suggests that the Y chromosome in *Tribolium* does not contain any male fertility genes. Also, all the progeny produced by virgin females mated with XX males were females, suggesting that a factor responsible for male sex determination is present on the Y chromosome of *Tribolium*. This was proposed by Shukla & Palli (2014) in a short communication. Our work supports these findings and their conclusion. However, given the difficulty involved in GMO regulation it is unlikely that this approach would be accepted.

The presence of *TcTra* above the threshold levels required to inhibit the activation of dosage compensation pathway in a few XX individuals may have helped them to escape zygotic death. However, the *Tctra* in these individuals may not have reached levels required to execute the splicing of *Tcdsx* pre-mRNA into the female mode, resulting in the production of masculinized females. This suggests there is a “Goldilocks” region of getting the gene expression (and the silencing) just right so that dosage-compensation is activated but splicing regulation fails. This could be optimised by timing of RNAi and/or dose of the dsRNA.

Injection lacks the scalability for a mass-rearing programme, it was important to ascertain the efficacy of the dsRNA in a more realistic setting. Thus, I attempted a feeding bioassay following methods described in the literature. dsRNA targeting the gene encoding the A subunit of vATPase had previously been shown to kill western corn rootworm (*Diabrotica virgifera*, LeConte) larvae in feeding bioassays (Baum et al., 2007), and here we observed that dsRNA targeting the *Tctra* female isoform failed to produce any phenotypic effects despite similar concentrations. However, our experimental design used adult insects as we wished to affect the sex ratio and induce pseudomaleness in their offspring. This difference may be due to idiosyncrasies of the gene being silenced or due to the adult not obtaining a sufficient dose for a phenotypic effect. This could be investigated by force feeding the female thus ensuring feeding of dsRNA.

Although the sex ratio was skewed towards more male and femaleness was silenced, as a method of silencing femaleness it was very inefficient. Levels of female contamination of mass-reared insects would be too high for a viable SIT programme.

The toxicity to females and the very low rate of pseudomale generation was modelled and described how underwhelming the mass-rearing method would be relying on this technology. For this approach to be viable the efficiency needs to be increased by orders of magnitude. As a method of silencing femaleness it seems inadequate compared to methods based on sex-specific splicing (Chapter 5). The benefit unique to this approach of converting females to pseudomales

occurs at too low a rate for it to affect the utility and remedy the number of females that would be released as part of an SIT programme. I believe the opportunity cost of optimising this approach is too high and the methods in Chapter 5 will provide a more fruitful foundation for a female-lethal based SIT programme in the Coleoptera.

The technology may be worth revisiting if methods for skewing the sex ratio towards more female could be achieved as this would cut costs in mass-rearing by reducing the number of unnecessary males produced. A ratio of 10:1 female to male would be more cost effective than 1:1 as complete fertilisation could still be achieved.

CHAPTER 7: THESIS SYNTHESIS



7.1 GENERAL DISCUSSION AND FUTURE CONSIDERATIONS

Heeding Malthus' warning (1798), scientists around the world continue to develop knowledge and technology to assist in maintaining global food production above that of global demand. Reducing crop losses to agricultural insect pests plays a pivotal role in meeting the increased demand associated with a growing global population.

Insects inflict crop losses of 15% annually (Maxmen, 2013) and warrant continued investment in developing strategies of sustainable control in contrast to historical petro-chemical intensive methods. The genetic toolbox developed during the last 15 years (Fraser, 2012; McGraw & O'Neil, 2013) combined with the historical sterile insect technique offer an environmentally friendly and scalable area-wide approach to insect pest management.

The plausibility of this approach has been vindicated by recent field trial successes of GM male mosquito releases producing a population suppression effect in the treatment area compared to the control (Harris *et al.*, 2012). This represents the first generation of RIDL technology which produces genetic control via tetracycline-repressible, bi-sex, late-acting dominant lethality (Phuc *et al.*, 2007).

This first generation technology is effective in suppressing *Aedes aegypti* (Harris *et al.*, 2012; Carvalho *et al.*, 2015), but not optimal, as it does not provide genetic sexing of the insects. When considering the system's utility against agricultural pest species, early-acting or embryo lethality would be preferred. This would minimise larval feeding damage to the crop, particularly important in crops with low economic thresholds for damage.

This thesis proposes a genetic SIT targeted against agricultural pests combining female-specific lethality with the paternal effect system, both conditional and regulated by the dietary supplement tetracycline. Hitherto this system has only been developed in the Diptera (Jin, 2011; Bilski, 2012) where final product lines are now in development by combining all three components into the third generation Oxitec product (1.1.9). This thesis sought to expand the applicability to the neglected Lepidoptera and Coleoptera insect orders.

The paternal effect system is dependent on identifying an appropriate endogenous promoter to drive sperm-specific expression prior to meiotic arrest (Jin, 2011; Bilski, 2012). The use of diamondback moth $\beta 2$ -*tubulin* as a suitable sperm-specific promoter for a paternal effect system was explored in Chapter 4. Due to its conserved nature, model gene status and having been implemented as a driver of the paternal effect system in Diptera the $\beta 2$ -*tubulin* gene in diamondback moth was pursued as the gene of choice. BLASTing the diamondback moth genome for $\beta 2$ -*tubulin* provided a putative gene. We confirmed its status as $\beta 2$ -*tubulin* by sequencing, alignment against dipteran and lepidopteran orthologues, RNA *in situ* hybridisation and RT-PCR of the testis compared to other tissues. The expression profile was shown to be similar to the dipteran orthologue. Comparison with that of *D. melanogaster* revealed a similar transcription pattern, with the gene transcribed in mid-early spermatocytes, with the transcript stable to late primary spermatocytes and persisting to early elongation. With the $\beta 2$ -*tubulin* gene being well-conserved between species, it is perhaps unsurprising that the transcription pattern and function of this structural gene remains similar in insects (Nielsen *et al.*, 2010).

This similarity to the paternal-effect system development in Diptera perhaps instilled overconfidence in the utility of $\beta 2$ -*tubulin* as a reliable go-to gene of the system. We expected the regulation of the gene to be similar to that of the previously investigated species and followed similar design principles in striving towards a paternal effect construct. Whereby the species' endogenous $\beta 2$ -*tubulin* promoter alone could drive expression of *Dmprot-DsRed2*. We were unable to achieve sperm-specific expression following this approach. A re-think occurred after 2 years of failure.

In hindsight the literature almost flaunts the relevant information with articles highlighting potential difficulties in using a lepidopteran $\beta 2$ -*tubulin* in this capacity. Raff *et al.* (1997) describe how in the moth *Heliothis virescens*: "*We obtained wild-type beta 2 tubulin-like levels of expression only with an intron-containing insert, suggesting that splicing may be important in normal beta 2 tubulin [β2-tubulin] expression.*"

We saw a similar phenomenon in diamondback moth where the intron-less transgene did not result in expression, whereas the inclusion of the first exon, intron and 18 nt of the second exon resulted in successful transcription of the transgene. The regulation of the gene does not seem as tightly conserved evolutionarily in non-Diptera (in the sense that the intron was required for expression). This difference could be explained by the absence (Diptera) and presence (Lepidoptera), of a testis-specific α -*tubulin* (with which the $\beta 2$ -*tubulin* pairs as a heterodimer in microtubule formation) described more fully in Chapter 4 (4.5). This finding should caution the application of learnings between orders or distant species, but lends credence to the development of the systems in model pest species prior to incorporation into more economically viable — from a commercial product development perspective — pests. This finding also illustrates that even a highly conserved protein family can participate in the adaptive process and respond to sexual selection (Nielsen *et al.*, 2010), particularly in the reproductive tissue-specific tubulin isoforms.

In a dipteran orthologue, activated gene expression is conferred by the core promoter in conjunction with appropriate transcriptional activator element(s). In *D. melanogaster* the promoter sequence responsible for tissue-specific gene activation is confined to a region of 80 bp, sufficient to drive germline specific expression in the testis. In addition, a 14-bp activator element (β 2UE1) is necessary for promoter specificity (Santel *et al.*, 2000). It would be useful to refine our knowledge of the diamondback moth $\beta 2$ -*tubulin* as we cannot make such precise statements. This could be achieved using standard transgenesis methods and/or by mutating putative crucial genetic regulatory components and examining the effect on transgene expression.

For diamondback moth the paternal effect system, through experimentation in this thesis, has progressed to the milestones of sperm-specific transcription and translation of a transgene. The endogenous $\beta 2$ -*tubulin* 3' UTR was necessary for sperm-specific translation as spermiogenesis proceeded. The $\beta 2$ -*tubulin*-tTAV was translated in the sperm and activated expression of *tetO*-*hsp70*-AmCyan providing the first instance of sperm-marking in a lepidopteran pest (Xu *et al.*, 2014). Further experiments should investigate the conditionality of the construct OX5196 by contrasting tTAV production (by proxy measurement of AmCyan in OX4026 cross insects) both on and off tetracycline. This would determine the feasibility of a product based on this system to express a protamine-nuclease under mass-rearing conditions. In parallel, mating experiments could be run to determine if AmCyan-positive sperm are detected in the spermathecae of mated females.

Most examples in the literature describe how the 3' UTR can interact with a specific protein that recognises its sequence to form an inhibitory closed loop via a repressor protein (or proteins) bridge to the 5' end (Jackson *et al.*, 2010). Thus translation is regulated by the 3' UTR protein interaction. For diamondback moth $\beta 2$ -*tubulin* I demonstrated the necessity of the 3' UTR in permitting translation. This suggests that translation regulation of transcripts involved in spermiogenesis after meiotic arrest may be activated rather than released from inhibition. This

could be investigated by examining the dependence of sperm-specific genes on their endogenous 3' UTR for translation. This contrasts to the translation of $\beta 2$ -tubulin in Diptera where the 3' UTR was unnecessary. This discrepancy could be a quirk of the Lepidoptera and their evolution of polymorphic sperm requiring more diverse regulation strategies during spermiogenesis.

Utilising this sperm-specific expression a nuclear-chaperone and nuclease will need to be expressed as a fusion protein. No progress was made in testing the viability of a diamondback moth chaperone-*Foql* fusion. This was due to the difficulty of locating the protamine homologue in the diamondback moth genome. An alternative gene could be sought with the desired DNA binding qualities. If no suitable candidates are found then the $\beta 2$ tubulin-*Dm-prot-Foql* transgene should be inserted testing its efficacy at sterilising the males while assessing mating-success and sperm delivery into the females using standard mating experiments (Ant, 2013; Harvey-Samuel, 2014). Alternatively, following RT-PCR to confirm a testes-specific expression profile of a protamine-like gene that has been identified in *Bombyx mori*, the gene could be fused to a fluorescent protein and tested using standard *piggyBac* insect transgenesis methods.

The translation profile of the $\beta 2$ -tubulin-tTAV may be too late for nuclease expression driven by tTAV, due to the lag time in translation of tTAV to *tetO-Foql* expression. By the time the nuclease is active in the sperm, the DNA may be coiled around protamines preventing nuclease access (person. comm., Koukidou). If the 3' UTR determines the translation timing of the transgene we may need to seek an alternative gene with earlier translation in diamondback moth spermatogenesis. The RISH provided a potential candidate that should be re-examined.

Tribolium castaneum is the model coleopteran pest species. Prior to my thesis this species had merely been transformed (Lorenzen *et al.*, 2002) and no work to develop genetic control in Coleoptera had been attempted. I started the work described in Chapter 5 by confirming the use of the endogenous α -tubulin-1 as a suitable promoter for reporting successful transgenesis (Siebert *et al.*, 2008). Marker reliability assays quantified ZsGreen to be more sensitive and

reliable compared to DsRed2 with a superior signal-to-noise ratio. Typical issues of auto-flourescence of ingested food in the gut and melanised cuticle were evident (Horn *et al.*, 2002). This thesis recommends *Tc- α -tubulin1-ZsGreen* as the transgenesis marker of choice when detecting successful insertion of a *piggyBac* payload. However, it was noted that, due to sperm-marking associated with the likely involvement of α -tubulin-1 in sperm axoneme formation, this module is not compatible with an additional parallel sperm marking system.

Following the development of the transgenesis marker module we focused on the female-specific expression module. We identified and sequenced both *Tc dsx* and *tra* but opted for the use of *tra* in our putative female-specific expression system. Rather than using a sex-specific promoter, the inclusion of the *tra* minigene was preferred with the females successfully splicing the male and shared introns leaving a correctly spliced transcript for DsRed2.

Despite this landmark success, no functional protein was detected via flourescence microscopy. This was an undesired feature of DsRed2 described throughout this thesis, which showed incompatability when used as a reporter with non-DsRed2 amino acids on the 5' end. It was suspected that this was caused by incorrect folding of the DsRed2, although this seems weakly plausible given the more central location of the chromophore at amino acid 66. The main issue with resolving the problem was the difficulty in interpreting DsRed2 descriptions from a chemistry-centric perspective provided by the literature (Pakhomov & Martymov, 2009). I lacked the language to form a coherent model and explanation for the failure of the DsRed2 in our constructs. Immunostaining techniques were inconclusive and could not provide data to test the DsRed2-failure hypotheses.

Chapter 5 finished with the design and construction of a female-specific lethal module conditional on tetracycline. Unfortunately the construction of this construct proved difficult and took several months to complete. This leaves the development of a paternal-effect system pending the testing of this construct. The splicing off-tetracycline could be assessed using conventional transgenesis methods with a comparison on and off tetracycline.

Given success we can proceed towards elucidating potential candidate genes for sperm-specific expression, perhaps moving away from (or towards?) the default choice of $\beta 2$ -tubulin given our learnings of Chapter 4. RNAseq transcriptome data could be used to determine genes with enhanced sperm-specific transcription by dissecting out and screening different tissues, similar to our approach in Chapter 4. Confirmation of appropriate genes for the paternal effect system could short-list candidates based on specificity and expression levels based on data produced by RT-PCR and RISH. Given our world-first application of the RISH to diamondback moth testes we are confident we could use the same method for *Tribolium*. In anticipation of transferring the system to other Coleoptera, the $\beta 2$ -tubulin gene and its regulation should be conserved within the Order.

Due to the long delay in construction of the final *Tribolium* construct described in Chapter 5, I used my time to investigate an alternative method of augmenting or adjusting gene expression in insects to be released as part of an SIT programme. In the work described in Chapter 6, I investigated the possible use of feeding dsRNA to *Tribolium* to induce RNAi, silencing genes associated with sex determination (an approach that I was keen to test since reading Fire *et al.*, 1998 as an undergraduate). This approach could be used to augment the mass-rearing of insects for SIT by either skewing the sex ratio towards more females during rearing to reduce costs and/or to silence femaleness in the release generation. The latter was attempted in work described in Chapter 6.

Although experimentally feasible, the use of *Tctra* dsRNA to skew sex ratios was a far inferior method of removing females from a release generation compared to a transgene based genetic-sexing system. Success was only achieved when injecting a range of different life-stages with dsRNA; a feeding bioassay showed no effect. This corroborates the series of papers of Shukla & Palli on sex determination in *Tribolium* (2012, 2013 & 2014). We found similar pseudomale generation rates and a negative effect on female survival, probably due to Tra protein being required for X-inactivation. Until the silencing efficiency is improved and oral delivery of dsRNA

to *Tribolium* is realised the technology should be shelved in favour of genetic-sexing transgenesis. Individual force-feeding of dsRNA to adult females may be informative. The current limitation of RNAi is the variability in efficacy. Future developments are likely to reduce the variability, perhaps through improved formulation (Zhang *et al.*, 2015), after all success through perseverance has been achieved in *Aedes* – why would *Tribolium* be different (Singh *et al.*, 2013)?

This thesis depended on insect transgenesis with *piggyBac*. Rather than depending on anecdotal evidence, I set out to elucidate the current state of the art and provide reliable estimates for microinjection survival and transformation efficiency across the full range of transformed insects in the literature and at Oxitec. Chapter 2 provided a useful meta-analysis, acknowledging publication bias and then proceeded to develop an online “app” for researchers to aid decision making; helping them to get the number of injections just right. Chapter 3 used a double-blind experiment to assess the inputs required for transient expression during insect transgenesis.

7.2 CONCLUSION

This thesis provides evidence of the viability of a paternal effect system outside of the Diptera. All the genetic components were not directly transferable from the Diptera to the Lepidoptera, with the use of the endogenous diamondback moth $\beta 2$ -*tubulin* non-coding DNA necessary to provide sperm-specific transcription and translation. Pending elucidation of appropriate components for DNA lysis this system is still several years from the design of a product line. The modular nature of the transgene design and the universality of the DNA imply that perhaps this paternal effect system may be transferable to other lepidopteran pest species, for example the fall armyworm (*Spodoptera frugiperda*, Smith) and corn earworm (*Helicoverpa zea*, Boddie) both of which exhibit high rates of insecticide resistance.

The lessons learned from the attempted implementation of the paternal effect system in the Lepidoptera will aid decision-making in developing the system in the Coleoptera beyond the

model species *Tribolium castaneum* into other important agricultural pest beetle species. The elucidation of a female-specific splicing system paves the way for future genetic control of any target Coleoptera.

It is hoped that this thesis contributes to the field of genetic pest management and the improvement of agricultural sustainability in the neglected insect orders. Beyond these two orders, this thesis also provided a systematic review and current state of the art of insect transgenesis using *piggyBac* in insects.

CHAPTER 8 MATERIALS AND METHODS

8.1 INSECT REARING

8.1.1 DIAMONDBACK MOTH

(After Martins *et al.*, 2012).

All laboratory experiments were performed within a temperature controlled room (25°C) with a 16:8 light:dark cycle. The diamondback moth wild-type background strain used for transformation and bioassays was provided by Syngenta plc. (Jealott's Hill, UK) in 2008 and originated in Vero Beach, Florida, USA.

Wild type stocks were reared in cages (350 mm x 230 mm x 230 mm). Adults were fed and watered using tetracycline sugar-water soaked cotton wool (Fisher, product code CTC-230-010S) in a pot. A few days after eclosion egg collection began; Parafilm strips (90 x 30 mm) were brushed with cabbage solution (cabbage leaves boiled in water; the water was used as an oviposition cue for the females). Strips were collected when sufficient eggs had been laid. Parafilm was placed into plastic pots (Fisher, product code FB74005). On emergence larvae were fed with Bioserve (USA) artificial beet armyworm diet (Frontier, product code F9221B). For tetracycline diet, chlortetracycline was added at a concentration of 100 µg/ml; non-tetracycline diet contained no chlortetracycline.

For large stock rearing of wild type Deli pots (Ambican, DP-DM12) could be used with a greater capacity. The lids had a 4 cm hole cut into them and cardboard (packaging2buy.co.uk) fitted into the hole allowing ventilation. Diet was poured directly into the Deli pot (2 cm deep) and allowed to cool.

8.1.2 PREPARATION OF SUGAR WATER; TETRACYCLINE / NON-TETRACYCLINE

Two litres of Milli-Q water were measured into a Duran bottle. Methyl 4 Hydroxybenzoate (Nipogen) (ELS, H6654-100G) was added and then the solution micro-waved on full power (800 W) for thirty seconds or until the Nipogen was dissolved. Sucrose (150 g) and tetracycline (Chlortetracycline Hydrochloride) (0.2 g) were added (ELS, 26430 -100G). A homogenous solution was created by thirty min of magnetic stirring and then stored at 4 °C for up to two weeks. The tetracycline was omitted for the non-tetracycline sugar-water version.

8.1.3 LARVAL DIET

Diamondback moth artificial diet ingredients were bought externally and prepared on-site (BioServe Biotechnologies, Ltd., USA). A 820 ml solution of water and 19.8 g of agar was brought to the boil using a microwave (800 W). The solution was allowed to cool below 65 °C and 161.8 g of the artificial diet powder was added. For tetracycline diet 0.1 g of chlortetracycline hydrochloride powder was added and homogenised with a blender.

8.1.4 *TRIBOLIUM CASTANEUM*

Our stock originated from the GA-1 strain provided by researchers at Kansas University, USA. This strain was the forbear to the sequenced GA-2 strain (Richards *et al.*, 2008).

Beetles were reared on a 95:5 mixture (by weight) of pre-sifted (710 µm) wheat flour and dried brewer's yeast (BTP Drewitt, UK). For egg collection type 405 "white" or "instant" flour (with yeast) was used. Diet was added to the container to a depth of three cm.

Insects were reared in cuboid plastic Tupperware containers (20 x 8 x 8 cm) with a 3 cm diameter hole cut into the lid. Containers were rendered insect proof by the gluing of a fine mesh over the hole thus allowing ventilation. The incubator was set to 32 °C with a relative humidity of between 20-40%. Data-loggers (Lascar electronics, UK) were used to ensure conditions remained within this range. At these temperatures the development time from egg to adult was expected to be about 30 days (Howe, 1965). The observed generation time was several days longer probably due to the low and variable humidity.

Adult beetles were removed and sub-cultured every week to prevent density-dependent disease, cannibalism and malnutrition. Females younger than 3 months old are more fecund than their older counterparts (Sokoloff, 1974). These were used for egg production and microinjection.

For stock populations, live adults were separated out from dead adults and exuviae by running string into the sieve and attaching to the side using blue-tack. Adults would instinctively climb up and out of the sieve leaving behind the dead, the infirm and the exuviae. If pupae were desired a similar process could be used to separate the adults from the pupae which could then be picked with soft-tweezers.

After Rietema, (1991), all life stages can be separated from nutritional media using an appropriately sized sieve:

- Adults and pupae are separated from earlier life stages using a 710 µm sieve (#25).
- Earlier instar larvae can be separated from later life stages using a 500 µm sieve (#35).
- Eggs can be collected from "instant" or tripled-sifted whole meal flour using a 300 µm sieve (#50).



Figure 140. A “normal” egg of *Tribolium castaneum* after Sokoloff (1972). Dimensions are $630 \pm 6 \mu\text{m}$ (S.E.) in length and $350 \pm 6 \mu\text{m}$ (S.E.) in width. Eggs were only visible following washing as the flour rendered them undetectable.

Experimentation revealed that some eggs were lost through the 300 μm sieve therefore the 250 μm sieve was preferred.

8.1.5 TRIBOLIUM CASTANEUM REARING ON A TETRACYCLINE-DOSED DIET

Five replicates (experimental unit at the level of the pot) contained five pseudo-replicates (number of organisms). Two different ages were assessed, 3-week-old adults or final-instar larvae, selected at random with five of each type per pot. Individuals were sieved out of flour and a number grid with Cartesian coordinates was assigned to the container floor, a random number generator was used to generate random coordinates. Insects were extracted with a fine brush and transferred onto flour with four grams of medium per pot (Campbell and Runnion (2003). Ambient conditions were kept constant at 26°C and 15% relative humidity in complete darkness. This set-up was chosen as it is considered “stressful”, as it was anticipated that stressful environments would amplify any differences between treatments. Adult survival was recorded weekly. Larvae were allowed to develop for a month and then examined weekly thereafter.

8.1.6 TESTIS ANATOMY AND DEVELOPMENT IN DIAMONDBACK MOTH

This was achieved by dissection in phosphate buffered saline (PBS) and imaged using light microscopy. Ten males of each life stage (10-day old larva, pre-pupa, pupa and adult) were dissected and examined with scalpel and fine tweezers. The most typical testes were photographed under standard lighting conditions. Testes diameters were measured using a graticule.

8.2. MOLECULAR METHODS

8.2.1 PREPARATION OF DNTPS

The four stocks solutions of ATP, GTP, CTP and TTP were thawed and 250 μl of each were added to 1.5 ml of Milli-Q water. Solutions were stored in 15 ml centrifuge tubes and frozen at -20 °C until required.

8.2.2 EXTRACTION OF GENOMIC DNA

8.2.2.1 FERMENTAS GENEJET KIT

Genomic DNA was extracted using Fermentas Genejet Kit (Fermentas International Inc., UK). Pupae or adults were preferred due to lower levels of body fat compared to larvae. Samples were placed into individual microfuge tubes and broken up in 180 µl of digestion solution and pulverized using a pestle. Proteinase K (20 µl) solution was added to the sample, vortexed and then incubated at 56 °C for 1-4 hrs. RNA was degraded over 10 min using 20 µl of RNase A solution. Lysis solution, 200 µl, was added. Samples were centrifuged if large amounts of insoluble materials were present (e.g. chitin exoskeleton). The supernatant was added to 400 µl of 50% ethanol solution and mixed. The sample was pipetted into a column and centrifuged for one min at 9000 rpm (Accuspin microfuge, Fisher Scientific). The column was washed and spun as appropriate using the kit recommendations. Finally the gDNA was eluted and stored at -80 °C.

8.2.2.2 PURELINK™ GENOMIC DNA KIT FROM INVITROGEN

Samples were placed into microcentrifuge tubes for homogenisation and digestion using a sterilised pestle and 20 µl of Proteinase K and 180 µl of Purelink™ Genomic Digestion Buffer. Samples were incubated overnight at 55 °C. The lysate was centrifuged at 13,000 rpm for 3 min at room temperature. The supernatant was transferred to a clean tube. 20 µl of RNase A was added to the lysate and mixed well by vortex and then incubated at room temperature for 2 min. A homogenous solution was created by adding 200 µl of Genomic Lysis / Binding Buffer and then vortexed. The lysate was mixed well with 200 µl of 96-100% ethanol and added to a Purelink™ spin column with a collection tube fitted. The column was spun at 10,000 g for 1 min. The tube was binned and replaced. Wash Buffer 1 was added at 500 µl to the column and centrifuged for 1 min at 10,000 g. This was repeated with Wash Buffer 2 but with a spin time of 3 min. DNA was eluted into a clean 1.5 ml centrifuge tube by addition of 200 µl Genomic Elution Buffer and then centrifuging at max speed for one min after a one min wait.

8.2.2.3 NON-ENZYMATIC GENOMIC DNA EXTRACTION

The legs were removed using dissection scissors. The leg was placed into solution A in a 0.2 ml PCR tube. It was then heated in a thermocycler at 99 °C for 30 min. Solution B was added and vortexed to neutralise ready for PCR applications (see 8.2.23 Leg PCR for genotyping for solution A and B details).

8.2.2.4 PCR BIO RAPID EXTRACT PCR KIT

Kit procured from PCR Biosystems Ltd. Protocol followed for 25 µl reaction. For the PCR 5µl of the extracted solution was used as template DNA.

8.2.2.5 RULING OUT CONTAMINATION OF OX4703 WITH OX4673

Two positive controls of 1:100 diluted OX4703 miniprep and OX4673 miniprep were included in the reactions. A negative control of just the primers and reaction ingredients were also included (PCR BIO Taq Buffer 4 µl, 10x bovine serum albumin 0.5 µl, primer F 0.2 µl, primer R 0.2 µl, BioTaq enzyme 0.2 µl, MilliQ H₂O 12.3 µl per sample; gDNA 2 µl or miniprep 0.5 µl). The thermocycler programme was determined by BIO PCR guidelines and length of DNA extension.

8.2.3 EXTRACTING RNA

8.2.3.1 QIAGEN RNEASY MINI KIT

Under the fume hood B-mercaptoethanol (B-ME) was added to the RLT buffer at 350 µl per sample, at a rate of 10 µl B-ME per 1 ml RLT buffer. The insects/samples were collected and stored at -80 °C prior to use. When required samples were homogenised by adding 600 µl of the prepared RLT buffer and the use of a pestle. Samples were centrifuged for 1 min at 3500 rpm to pellet the chitin. The supernatant was transferred to a Qiashredder column (Qiagen Ltd., Crawley) using a P1000 set to 300 µl. One volume of ethanol (70%) was added to each sample and mixed using pipette to facilitate the precipitation of RNA. RNeasy spin columns were used to filter the solution with the RNA finally eluted at the end of the process. RNA was stored at -80 °C until required.

8.2.3.2 RNA EXTRACTION FROM TRIBOLIUM CASTANEUM FOR DOUBLESEX RT-PCR

Tribolium pupae were sexed and placed individually into Eppendorf Standard Micro Test Tubes 3810X. Proteins and enzymes were deactivated by the addition of 50 µl of phenol. The insect samples were homogenised in a fume cupboard. The homogenate was incubated for 5 min at room temperature. The homogenate was then centrifuged at 12,000 g for 10 min at 4°C. The supernatant was transferred to a fresh tube. Chloroform was added and then vortexed briefly. After 10 min of room temperature incubation the tube was centrifuged for 10 min at 4°C. Separate layers formed with the colourless layer being transferred to a fresh tube and washed in isopropanol. The isopropanol was separated by centrifugation, the pellet was washed again with 75% ethanol. The ethanol was removed by pipette and air drying briefly. The RNA pellet was dissolved in 30-50 µl of fresh Milli-Q water.

As an additional quality control measure to check that RNase contamination had not occurred, 1 µl of each independent RNA solution was run on an agarose gel. The formation of one defined band of ribosomal RNA was expected.

8.2.3.3 RNA EXTRACTION AND RT-PCR TO ASSESS TRANSCRIPTION

8.2.3.3.1 ENDOGENOUS B2-TUBULIN TRANSCRIPTION IN THE TESTES

RNA was extracted using Qiagen RNeasy (Qiagen Ltd., Crawley) mini kit. The RNA was run in a gel to check for degradation. It was quantified using the Nanodrop spectrophotometer (Thermo Fisher Scientific, Inc., USA) and then run in a one-step RT-PCR for primers targeting the transgene of interest.

8.2.4 REVERSE TRANSCRIPTASE PCR USING REVERTAID

Following extraction, RNA was converted to cDNA using RevertAid (Fermentas International Inc., UK). The cDNA was amplified by PCR to make multiple copies. The combined procedure is referred to as reverse transcriptase PCR (RT-PCR).

Samples were removed from freezer and thawed. Ten µl of RNA was added to 1 µl of DNase 1 incubation buffer, 0.5 µl of Ribolock and 0.5 µl of DNase 1.DNA was removed from the sample by incubating at 37 °C for 20 min. Heat deactivation was achieved by adding 0.3 µl EDTA and heating at 75 °C for 10 min. The RNA was run on a gel to check for quality and presence. RNA

was quantified at 1:50 dilution (in water) from lowest concentration to highest, judged from the gel results, using the GeneQuant 2 or NanoDrop.

After quantification the appropriate amount of RNA for each sample were added to Milli-Q water to make up a total volume of 11.5 μ l at a concentration of 0.5 μ g RNA per 20 μ l. One sample was duplicated and used as a no reverse transcriptase control. To the RNA/water 1 μ l Oligo(dT) (10 μ M) were added. Oligo(dT) bind to the base pairs of the poly(A) tail on the 3' end of eukaryotic mRNA. Each sample was warmed at 65 °C for 5 min.

The PCR programme continued after the addition of 4 μ l 5x Reaction buffer, 0.5 μ l Ribolock, 2 μ l 10 mM dNTP mix and 1 μ l of RevertAid (except for the no RT control). The programme continued converting the RNA into cDNA. The cDNA was stored at -80 °C or used immediately for a general PCR to check confirm presence of a known mRNA (now cDNA) sequence in a given sample by use of appropriate primers. A control of primers for endogenous or housekeeping genes was used. Negative control of water and positive control of the cDNA of interest were included if possible.

8.2.5 QUANTIFICATION OF DNA AND RNA USING GENEQANT 2

The aromatic rings of the bases found in DNA and RNA absorb ultraviolet light with an absorption maximum at 260 nm (Clark & Pazdernik, 2012). If dissolved in solution the proportion of UV transmitting through the solution, when compared to a calibration control (just the diluent e.g. Milli-Q water), can be used to deduce the concentration of nucleic acid on comparison to a standard curve. This procedure is automated using a spectrophotometer.

RNA and DNA content were measured using the GeneQuant 2 spectrophotometer (GE Healthcare Ltd., UK). Samples were diluted to 1:50 and then the nucleic acid content quantified relative to a control of Milli-Q water. The concentration was read twice for each sample, these pseudoreplicates were averaged away to produce a concentration for each sample.

Towards the end of this thesis quantification was achieved using recommended guidelines on a NanoDrop spectrophotometer (NanoDrop products, USA).

8.2.6 POLYMERASE CHAIN REACTION (PCR)

"PCR requires template DNA, primers, a polymerase, nucleotides and a thermocycler to alternate the temperature of the reaction" (Clark & Pazdernik, 2012).

DNA sample combined with other ingredients to create a PCR reaction mix containing: 0.2 mM dNTPs, 0.2 μ M forward primer, 0.2 μ M reverse primer, approximately 100-300 ng of gDNA, 1x reaction buffer and the recommended amount of DNA polymerase in 25 or 50 μ l final volume. Milli-Q water is used in PCR reaction mixes. These default details may have been adjusted if amplification was proving to be difficult.

Reaction mix was added to PCR tubes and placed in PCR machine Biometra T3000 or Biometra TGradient thermocycler, Biometra, Germany) and run using an appropriate programme following manufacturers recommendations.

DNA polymerases included Herculase (Agilent Technologies Inc., USA) and Phusion (Fermentas International Inc., UK). Half-way through the thesis PCR BIO Taq DNA polymerase (PCR Biosystems Ltd., UK) was used. It was preferred as reaction buffer and dNTPs were pre-mixed.

Thermocycler programmes followed recommended guidelines for the product in use adjusted for the size of the product and primers used as appropriate.

8.2.7 PRIMER CHOICE

The gDNA sequence of the appropriate gene was copied into VectorNTI (Life Technologies Corporation) where exons and introns were marked by comparison with mRNA sequence and knowledge of typical splicing rules. Following annotation the primers were designed to target a specific region and have a software-predicted annealing temperature of 56°C and ordered from Life Tech® (formerly Invitrogen).

For the RISH primer design the coding region was assessed for whether it was in the open reading frame. Several of the genes had lots of short exons, described diagrammatically below. Accordingly several RT-PCR reactions (different reverse primers) were required per gene to elucidate any splice variants and to ensure that our future probe design would not miss the testis specific transcript of interest, furthermore different splice variants can exist in the soma compared to the testis.

The primer name contains information regarding what gene from HongFei Gong's list it is associated with (e.g. Nr6 is number six in the list except where the gene has been confirmed as *β2-tubulin*) and the relevant exon (e.g. E2 is exon two (assuming we have all the mRNA)) and the primer direction (forward or reverse). The number of reactions and associated exons with the reverse primers was chosen in an arbitrary fashion.

8.2.8 SEQUENCING

For sequencing DNA samples were sent with appropriate primers to GATC biotech (London). Sequences were examined using Vector NTI (Life Technologies Corporation).

8.2.9 SEQUENCE ALIGNMENT

Alignment of two or more nucleic acid sequences was achieved using AlignX functionality in Vector NTI with default parameters (Life Technologies Corporation).

8.2.10 AGAROSE GEL ELECTROPHORESIS

See Brody & Kern, 2004.

8.2.12 TRANSFORMATION USING XL-10 GOLD ULTRACOMPETENT CELLS

One tube of XL10-Gold Ultracompetent Cells, and XL10-Gold β-mercaptoethanol mix were allowed to thaw on ice. After adding 4μl of β-mercaptoethanol mix to the 150μl of competent cells the reagents were mixed gently and left on ice for 10 min with intermittent swirling.

Chilled 14 ml round bottomed polypropylene falcon tubes were loaded with 2 μ l of each ligation. Cells were divided evenly between tubes ensuring placement of cells on DNA droplet. Cells were mixed gently and placed on ice for 30 min. The cells were then plunged into a 42 °C pre-heated water bath for 30 seconds and promptly returned to ice for two min. Post heat shock 200 μ l of transformation medium was added to each tube. The tubes were incubated at 37 °C with vigorous shaking (~250 rpm) for one hr. The cells were then plated on antibiotic plates under sterile conditions with the addition of 5 ml of IPTG per plate. Plates were incubated at 37 °C overnight.

8.2.13 PICKING COLONIES FOR PCR SCREENING AND PREPARING CULTURES FOR MINIPREPS

LB broth was prepared and 100 μ l added to each well of a sterile 96-well culture microplate, for the number of colonies to be screened. A PCR reaction was required for the screening process to confirm a colony contained the desired sequence of DNA. A general PCR reaction was mixed with primers from the CloneJET PCR cloning kit (Thermo Fisher Scientific Inc.) and placed into a PCR reaction tube. A random colony was chosen using a random co-ordinate system, a pipette tip was used to pick the colony and it was gently submerged into the appropriate PCR reaction tube. The corresponding well was then dipped into by the pipette with a slightly more vigorous action to ensure bacteria entered the broth. When the required number of colonies was picked (usually 20 per sample plate) both plates were sealed. The PCR reaction tubes were placed into a thermo-cycler and appropriate extension time used for the product. The 96-well culture plate was shaken vigorously (~200 rpm) at 37 °C for an hr.

PCR products were run on a 0.8% agarose gel at 120 V for 20 min. The colonies testing positive for the expected band were used to inoculate minipreps (e.g. two positive colonies per sample), adding 20 μ l of culture to a labelled 15 ml tube containing 3 ml of LB Broth (with ampicillin). Cultures were incubated overnight at 37 °C and 250 rpm.

8.2.14 ISOLATION OF PLASMID DNA FROM E. COLI TO PRODUCE A MINIPREP

Following overnight growth samples of E. coli saturated broth that tested positive in screening were poured into microfuge tubes and spun at 13,000 rpm for 30 seconds. The supernatant was disposed of, the pellet re-suspended by vortexing and addition of 250 μ l of Resuspension Solution.

The GeneJET plasmid kit (Thermo Fisher Scientific Inc.) was used and instructions followed therein to produce a miniprep.

8.2.15 5'RACE USING SMARTER RACE CDNA AMPLIFICATION KIT

The kit was procured from Clontech Laboratories, Inc.

The RNA at 1 μ g / μ l were added to a microcentrifuge tube with 1 μ l of 5' CDS Primer A. The volume was made up to 3.75 μ l using sterile water. The tube was incubated in thermocycler at 72 °C for 3 min, 42 °C for 2 min. The sample was added to 2 μ l of 5X first-strand buffer, 1 μ l DTT (20mM) and 1 μ l dNTP mix. The contents were mixed using a pipette. One μ l of the SMARTer IIA

oligo were added per reaction. At room temperature I added 4.0 μ l of the buffer mix, 0.25 μ l RNase inhibitor and 1.0 μ l SMARTScribe Reverse Transcriptase (100U). The tube was incubated using a thermocycler at 42 °C for 90 min and 70 °C at 10 min. The sample was diluted with water to an appropriate final concentration, 100 μ l.

RACE carried out using universal primers and appropriate designed specific primers in a standard PCR.

8.2.16 QIAGEN MINELUTE PCR REACTION CLEAN UP

Following up from a digestion using a restriction enzyme and buffer combination the DNA product was cleaned. The ERC buffer, 300 μ l, was added to 100 μ l of the enzymatic reaction mix. The sample was applied to a MinElute column in a 2 ml collection tube and spun at 13,000 rpm for 1 min. The flow-through was discarded and the column washed with 0.75 ml PE buffer and centrifuged at 13,000 rpm for 1 min. Flow-through discarded and previous step repeated. The DNA was eluted into a clean tube by adding 10 μ l of EB buffer to the membrane and centrifuging at 13,000 rpm for 1 min.

8.2.17 PURIFICATION AND EXTRACTION OF DNA USING AN AGAROSE GEL

The gel was placed on a transilluminator and the band cut away from the gel using a scalpel. The band was placed in a microcentrifuge tube with 3 volumes of QG buffer to one volume gel. The gel was dissolved at 50 °C for 10 min. The sample was added to a QIAquick column and centrifuge for 1 min at 13,000 rpm and again with 0.5 ml of buffer QG at 13,000 rpm for 1 min. The column was then washed with 0.75 ml of PE buffer and centrifuged at 13,000 rpm for 1 min. To elute the DNA, 50 μ l of EB buffer was added to the column membrane and centrifuged in a 1.5 ml microcentrifuge tube for 1 min.

8.2.18 DNA PURIFICATION USING QIAGEN QIAFILTER ENDOFREE PLASMID MAXI KIT

Kit procured and the protocol was followed.

8.2.19 HIGH FIDELITY PCR USING Q5 POLYMERASE (M0491)

The kit was procured from New England Biolabs., the protocol was followed.

8.2.20 ONE-STEP RT-PCR USING SUPERSCRIPT III WITH PLATINUM TAQ

The kit was procured from Thermo Fisher Scientific Inc., the protocol was followed.

The PCR was set up in a reaction tube with the following ingredients: 10 μ l of 2 x reaction mix, 1 μ l of template RNA, 0.4 μ l of the forward and reverse primers, 0.8 μ l of SuperScript III RT / Platinum Taq mix made up to 20 μ l using nuclease free water.

Recommended temperatures of 55 °C for the cDNA synthesis step and the annealing step for PCR were used. An extension of 68 °C for 1 min was applied with the number of cycles totalling forty. A final five min extension period at 68 °C was included.

If semi-quantitative approach was desired then the thermocycle would be paused at the appropriate cycle number at 68 °C. The PCR product required volume (3-4 µl) would be extracted using a multi-channel pipette. Products run on the same gel for comparison.

8.2.21 *IN VITRO* TRANSCRIPTION OF RNA USING MEGASHORTSCRIPT KIT (T7 PROMOTER)

The kit was procured from Thermo Fisher Scientific Inc. the protocol was followed.

The reaction was assembled in an RNase free PCR reaction tube. Eight µl of RNTP solution (1:1:1:1 of ATP, CTP, GTP and UTP at 75 mM prior to mixing). The template DNA was added 300-500 ng with 2 µl of the reaction buffer and 2 µl of the enzyme mix. The tube was mixed and centrifuged prior to 4 hrs incubation at 37 °C in a thermocycler. The product was resolved on a gel. The RNA was precipitated and then resuspended in injection buffer prior to pupal injection in *Tribolium*.

8.2.22 WESTERN BLOT OF DSRED

Protein samples were prepared by pestle and mortar. The supernatant was mixed with 4x reducing loading buffer in the fume hood (Laemmli Sample Buffer and beta-mercaptoethanol at 9:1 ratio). Fifteen µl of sample was mixed with 5 µl of the loading buffer in a 2 ml microcentrifuge tube. Samples were spun down and locked prior to 5 min at 95 °C. Samples were run on a 4-15% Mini-Protean TGX gel (BioRad, Cat # 456-1096). Gels were run at 250 V constant for twenty min in 1 x SDS/glycine/Tris buffer.

Proteins were transferred onto a nitrocellulose membrane. Membrane was blocked with 1 x Pierce Clear Milk Blocking Buffer for 1 hr on a rocking platform. The membrane was incubated with primary antibodies for 1 hr rocking at room temperature. The membrane was washed of antibody four times at 20 min in TBST, then incubated with secondary antibody and washed. The membrane was developed with ECL clarity substrate prior to imaging.

8.2.23 LEG PCR FOR GENOTYPING

To develop homozygous lines or set up female by pseudo-male (XX-XX) crosses for Chapter 6, one had to determine the genotype of an insect without destroying it. This was achieved using leg PCR (Carvalho *et al.*, 2009) where a leg was removed and placed into 25 µl of 10 µL of Protease K (MP Biomedicals, Solon, OH, USA) at 400 µg/ml buffered in Solution A (10 mM Tris-Cl at pH 8.2, 1 mM EDTA, 25 mM NaCl). It was heat treated at 95°C for 30 min. Five µl of Solution B were added to neutralise the solution (turning it yellow when mixed), ready for use in a genotyping PCR (take 2 µl).

8.2.23.2 LEG PCR TO COMPARE SEX GENOTYPE TO PHENOTYPE

Prior to assessment of injection survivors it was prudent to assess non-destructive genetic sexing methodology by leg PCR.

Survivors were screened and sexed by phenotype at pupae or adulthood. Genetic sexing using sex specific primers in a multiplex PCR was conducted to test for RNAi of *Tctra*. A non-enzymatic method was used to extract gDNA from the leg removed from control insects to test whether insects could be genotyped for sex non-destructively prior to mating with wild type females (Figure 141).

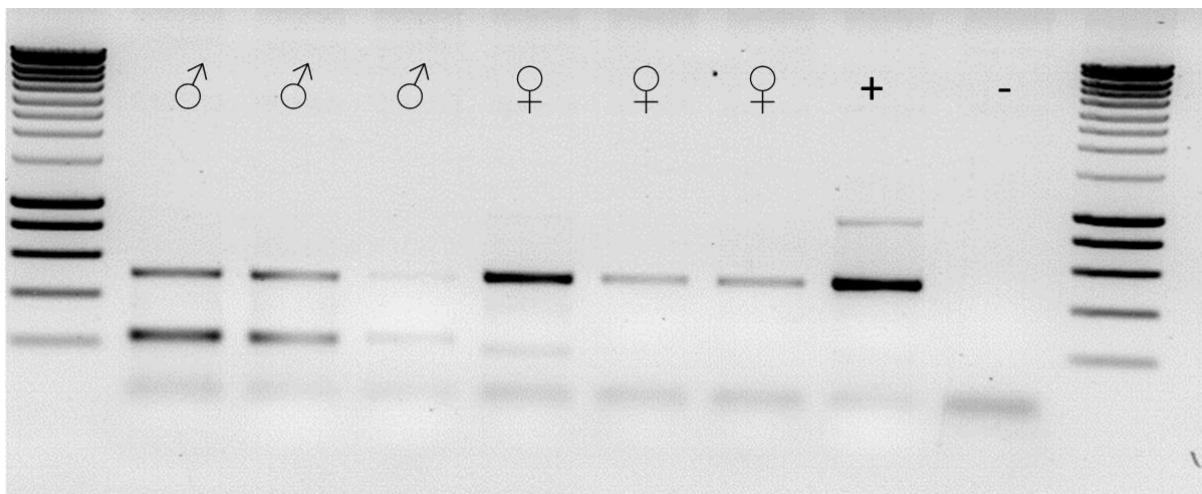


Figure 141. Assessing leg-PCR methodology performance in determining genetic-sex against *Tribolium* of a known sex. The middle left leg of males and right of females was removed with dissection scissors and heated to 99 °C for 30 min in solution A. The basic solution was neutralised with solution B prior to PCR. Primers TD3372, 3373 and 3374 were used in a multiplex PCR with both sexes having an expected band at ~ 500 bp and the males with a specific 230 bp band. Adults were sexed phenotypically prior to leg removal, by front leg male dark patch examination. A ~500 bp band was detected in males and females, a ~230 bp band was found exclusively in PCRs with male DNA as expected. The positive control was OX4700Z that had previously been identified as female. The primers were not contaminated as evidenced by the negative control. Additional non-specific banding is also detected and expected in some of the samples. The left most sample of each sex was from the femur down, the second sample the tibia and tarsus and the third, the tarsus and claw. The sex of individuals at adulthood was identical to the sexing at pupae suggesting both methods are robust. Male three was very weak but the bands can be seen. Assuming removal of the tibia downwards does not affect survival and reproduction this may be preferred as it is easier to dissect than the whole of the leg due to the thickness of the femur. The primers were developed by Lagisz *et al.* (2010).

8.2.24 MULTIPLE PCR IN *TRIBOLIUM* FOR SEXING GDNA

Multiplex PCR reactions were set up with the three primers (TD3372, TD3373, TD3374) and the four gDNA samples plus a negative control without any gDNA (reaction mix, primers and polymerase). It was expected that a 500 bp band would be detected in both males and females, whereas a 230 bp band would only be present in PCRs with male DNA. OX4700A and OX4700Z gDNA were derived from females; OX4700E and the wild type sample were male. As shown in Figure 151, there were additional bands (~900 bp), noted by Lagisz *et al.*, (2010) but these can be ignored as our interest lies with the two informative bands.

8.2.25 RNA *IN SITU* HYBRIDISATION IN DIAMONDBACK MOTH

Prior to the RISH candidate genes were assessed for testes-specific expression using RT-PCR on different target tissues in diamondback moth (Figure 142). Given amplification in the testes containing tissues only probe design proceeded (Figure 143).

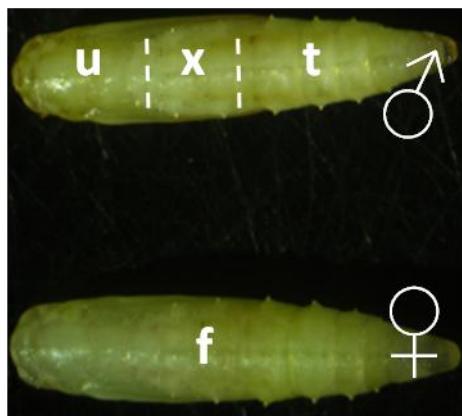


Figure 142. Pupae are shown here but the same principle applies for larva and adult dissections of testes and non-testes regions for RT-PCR. The whole female insect is used, the male is approximately trisected into the upper (u - anterior region, no-testes), testes (t – testes containing posterior region) and a buffer zone to be discarded from analysis (x – no testes, discarded to mitigate the risk of contamination).

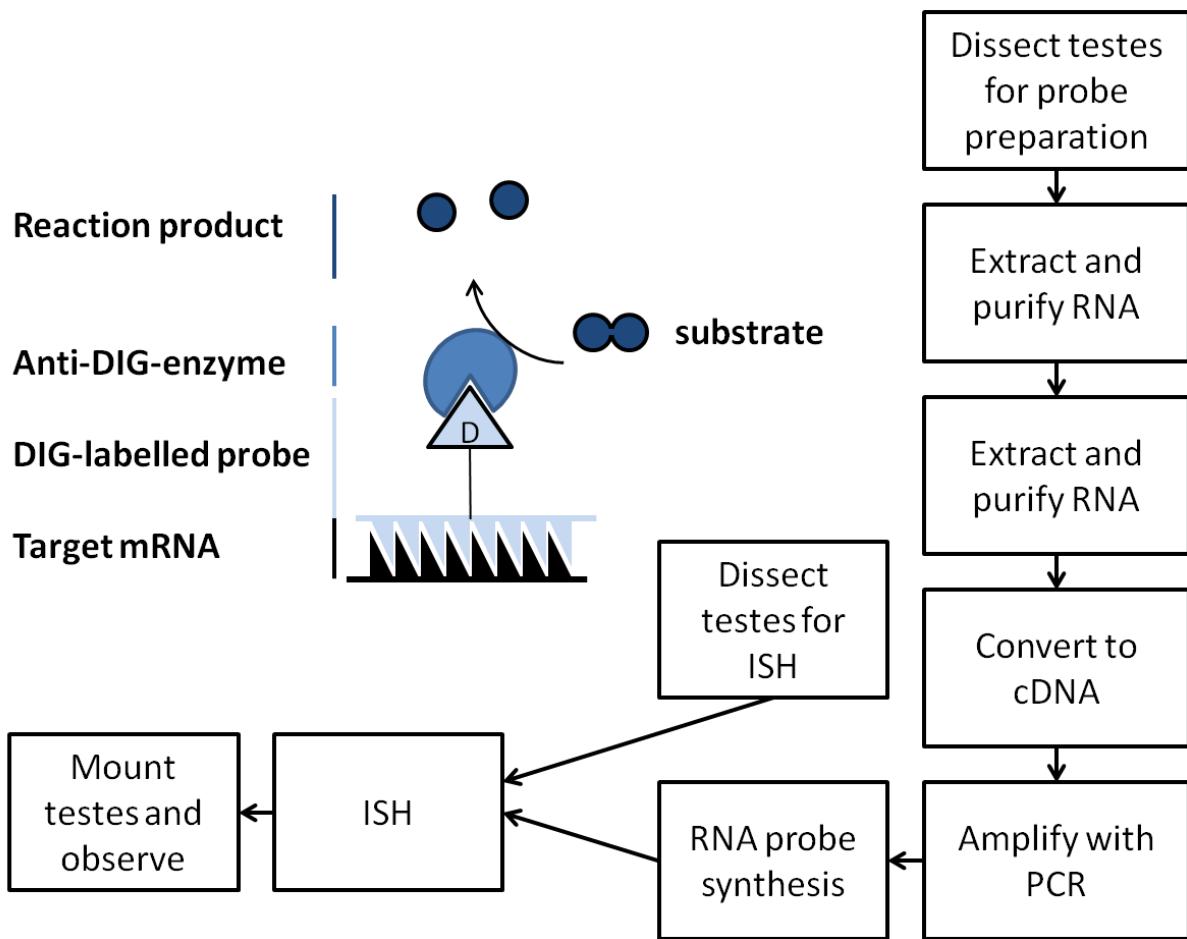


Figure 143. RNA *in situ* hybridisation allows visualisation of gene expression in cells within the testes for a target mRNA. The key steps are outlined in the boxes on the right. The most difficult step is developing appropriate primers to use for the "Amplify with PCR step". This requires gDNA and mRNA sequence knowledge for the target gene. The outcome is cells that are positive for a target mRNA will produce the reaction product which can be viewed under a microscope.

8.2.25.1 PROBE DESIGN

The linear PCR products described above were used as the template to produce anti-sense RNA probes following the methodology of Morris *et al.*, (2009). Sense probes were also produced to act as a negative control for an mRNA probe with equivalent GC content and can readily be prepared by incorporating an RNA polymerase promoter site in the 5' PCR primer. The primers used were based on the RT-PCR discussed in Chapter 4. Primers were designed against the gene of interest with the 3' primer including an RNA polymerase recognition site. Primer sites including introns were preferred so differences in length between cDNA and gDNA are detectable.

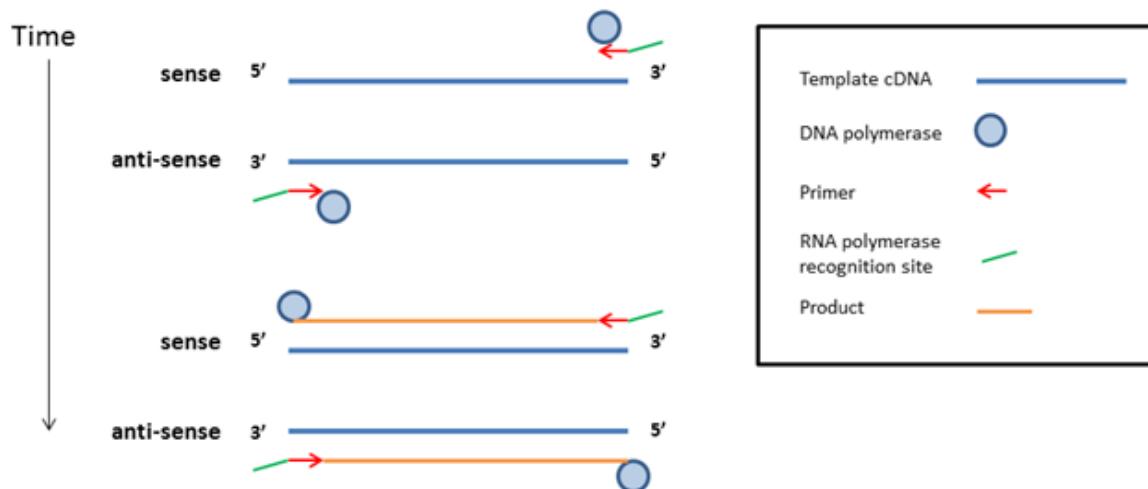


Figure 144. A diagram to aid the reader in following primer design and probe production. Primers (red arrows) were designed for the candidate genes with VectorNTI and incorporated a sequence of RNA polymerase consensus sequence to facilitate conversion of the sequence into complementary mRNA when the probe is made. Here both primers contain the RNA polymerase recognition site (green line) but each would have a different version depending on whether it was sense (T3 promoter) or anti-sense (T7 promoter). The diagram shows how the DNA version of the gene is amplified with RNA polymerase recognition site on the 3' end of the product.

The diagram above gives a pictorial representation of the amplification of the candidate gene with a 3' RNA polymerase site facilitating the use of RNA polymerase to convert the sequence from DNA to RNA required for the probe.

8.2.25.2 PRIMER CONSIDERATIONS

The prefix includes a four digit identification number followed by the putative gene name and the type of RNA polymerase recognition site attached (T3 consensus promoter for the negative or sense and T7 consensus promoter for the positive or anti-sense). The aforementioned recognition site was protected with a four nucleotide sequence of 'cgtc' to ensure good transcription from the 5' end of the 3' primer and to avoid errors in the consensus region during primer construction. For the T7 we took the complementary of the selected primer.

Table 44. RNA polymerases and the consensus sequences from thermoscientificbio.com.

| RNA polymerase | Consensus |
|----------------|-------------------------|
| T3 | AATTAACCCTCACTAAAGGGAGA |
| T7 | TAATACGACTCACTATAGGGAGA |

These polymerases are very similar in sequence although are described as having strict specificity for their double-stranded DNA promoters. To err on the side of caution and avoid any leaky transcription it was decided to set up two reactions per gene for the sense and anti-sense probe amplification. For the T7 associated primer sequences we took the complementary sequence as denoted by upper case letters (Appendix 9.3).

8.2.25.3 REACTIONS

This left us with five genes and two different reactions per gene to amplify mRNA suitable for *in situ* hybridisation probe construction.

Table 45. The successful primer pair reactions that gave rise to strong bands in the RT-PCR mentioned previously were used in primer design for another series of reactions to create a template for RNA polymerase to work on and create the appropriate RNA probe for the intended *in situ* hybridisation application.

| Gene (top twenty testes transcription list and putative homologue name) | Evidenced by RT-PCR primer reaction | RNA polymerase consensus promoter containing primer | Primer partner | Reaction label | Expected band size approx. (bp) |
|--|--|--|-----------------------|-----------------------|--|
| #1 CG8121 | iii | 3663-CG8121-T3 | 3453 | 1 | 734 |
| | iii | 3664-CG8121-T7 | 3452 | 2 | 734 |
| #5 CG9222 | ii | 3665-CG9222-T3 | 3455 | 3 | 706 |
| | ii | 3666-CG9222-T7 | 3454 | 4 | 706 |
| #6 | i | 3667-DBM-No6-T3 | 3634 | 5 | 466 |
| | i | 3668- DBM-No6-T7 | 3633 | 6 | 466 |
| #17 | i | 3669-DBM-No17-T3 | 3640 | 7 | 472 |
| | i | 3670-DBM-No17-T7 | 3639 | 8 | 472 |
| #4 β2 tubulin | i | 3671-DBM-B2T-T3 | 3638 | 9 | 574 |
| | i | 3672-DBM-B2T-T7 | 3637 | 10 | 574 |

The PCR reactions were carried out with Q5 polymerase due to its higher fidelity. Any inaccuracies in sequence for the RNA polymerase consensus sequence would be unacceptable. To avoid repeating a one step RT-PCR the short cut of using the RT-PCR two step products as the PCR template was used. One µl of the PCR product from the appropriate reactions described above.

The PCR product was checked by gel electrophoresis.

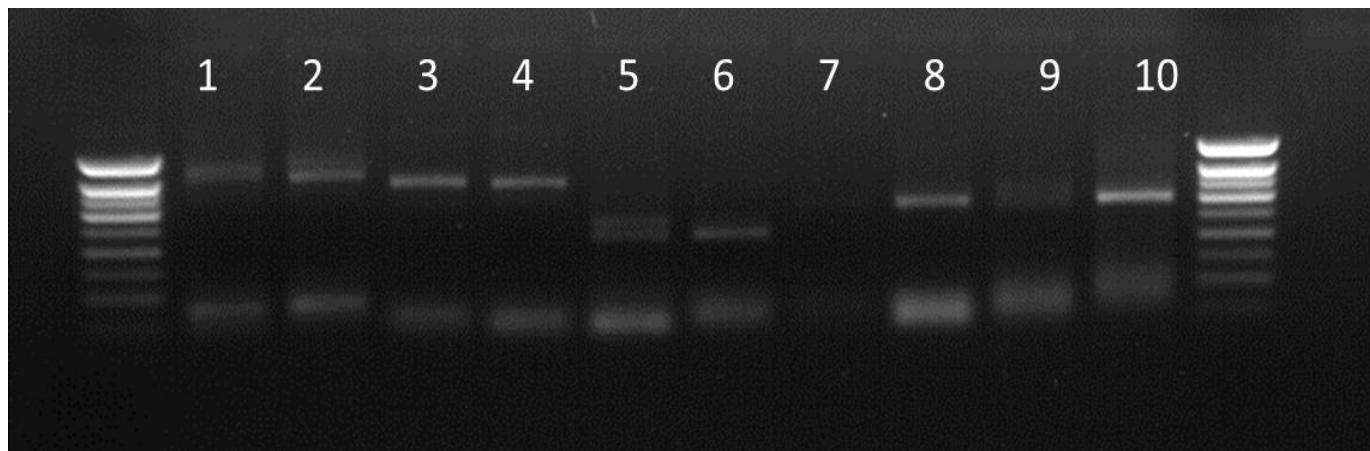


Figure 145. High fidelity Q5 polymerase PCR for diamondback moth testes specific genes to amplify for probe creation. Reactions are detailed in Table 45. All reactions amplified successfully except for reaction 7 which shows a very weak ghost band. All reactions were underwhelming in the concentration of the final product. Bands are of the expected size with some reactions showing unusual double bands not seen previously (reactions 2, 5 and 9). Surrounding lanes provide DNA size standards.

The results were underwhelming; inadequate concentrations of the product for those that did amplify and no product (perhaps a ghost band) for reaction seven. Use of the PCR product from the previous assessment of transcript type and location may be the problem with evidence that primer carry over is causing double band formation in reactions two, five and nine. The original RT-PCRs were repeated with the one-step method and then the PCR products used in the reactions described above to overcome this problem.

8.2.25.4 A DIFFERENT APPROACH: ONE-STEP RT-PCR

A one step RT-PCR using Superscript III with Platinum Taq (Invitrogen, Life technologies, www.lifetechnologies.com) was conducted following the default protocol provided. The products were run on a 1.2% agarose gel for 30 min at 120 V. The amplification was superior to that achieved with the two-step method. All the bands produced were of the expected size except for those targeting gene #17 (Figure 146).

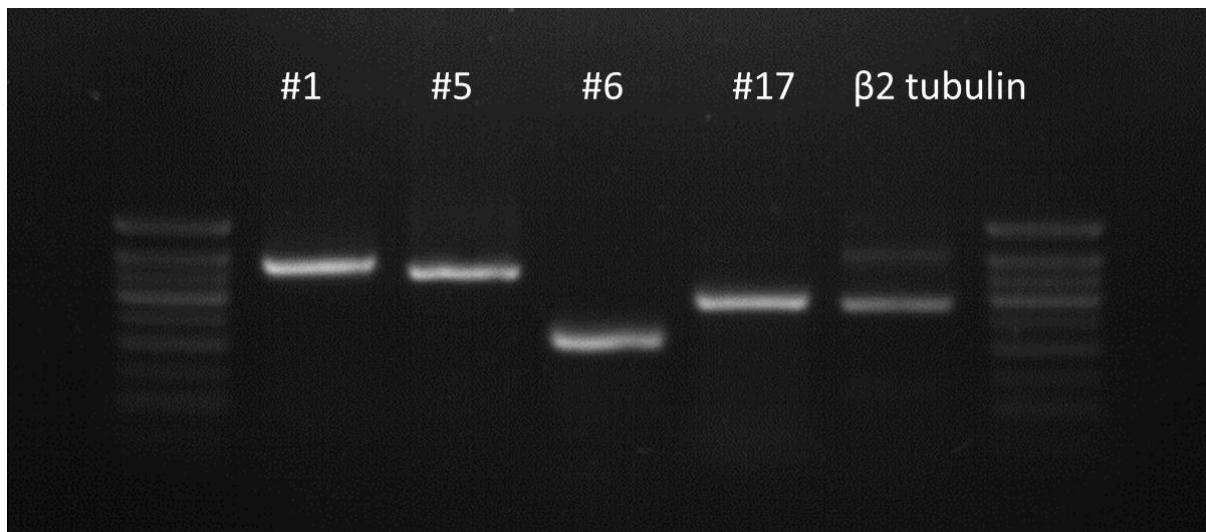


Figure 146. One-step RT-PCR of diamondback moth male carcass with testes specific gene labels provided (names based on position of genes in top transcription testes specific genes list). Primers for the reactions were, from left to right: 3452 & 3453, 3454 & 3455, 3633 & 3634, 3639 & 3640 and 3637 & 3638. Three μ l of product and 2 μ l of 6x loading dye were mixed before loading onto a 1.2% agarose gel. The gel was run at 120 V for 30 min. All bands were clear and of the expected size: #1 734 bp, #5 706 bp, #6 466 bp, #17 472 bp and β 2-tubulin 574 bp; except for #17.

The predicted band size for #17 was 472 bp, it appears to be between 500 and 600 bp when compared to the ladder; it is about 100 bp larger than expected.

Cloning and sequencing confirmed the presence of a previously unidentified exon of 105 bp in size from 891-996 base position in the gDNA sequence for this particular gene. This finding updated our records and cleared the gene for probe development. Five silent point mutations were present. With the additional sequence information it may improve results from a BLAST search of an orthologous gene.

The PCR products (except for #17 which was conducted after sequencing) were amplified with the primers described in Table 45. The Q5 high fidelity polymerase was used on the one-step RT-PCR products with the appropriate primers, two reactions per gene producing template for sense and anti-sense probes (Figure 147).

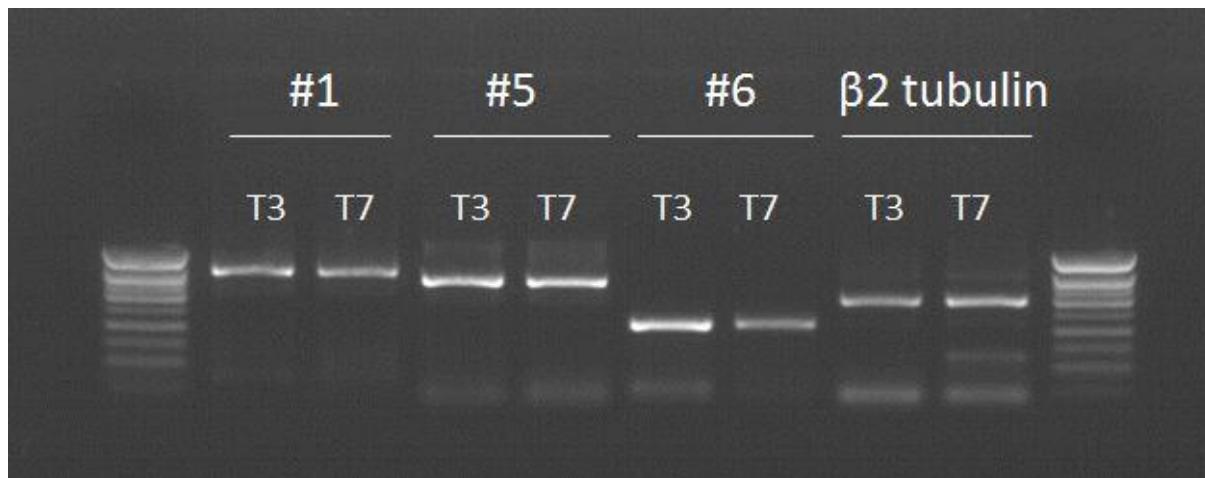


Figure 147. High fidelity Q5 polymerase PCR for diamondback moth testes specific genes to amplify for probe creation. Reactions are detailed in Table 45. Bands are of the expected size with some reactions showing unusual double bands not seen previously.

The concentration is improved and adequate. The β 2-tubulin T7 reaction produced a double band requiring gel purification step to avoid the extra band. The other products required column purification before RNA probe construction.

The same procedure was conducted on gene #17 in preparation for the final mRNA probe production.

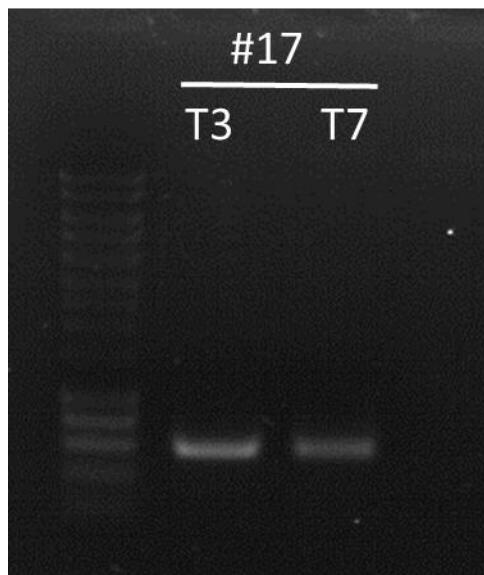


Figure 148. High fidelity Q5 polymerase PCR for diamondback moth testes specific genes to amplify for probe creation. Reactions are detailed in Table 45. Bands are at the expected size of approximately 577 bp following the discovery of the additional exon. Five μ l of smart ladder was used with 3 μ l of PCR product and 2 μ l of x6 loading dye in the adjacent wells.

The product was present with a clean band of the expected size; the image taken was faint, likely due to the gel being made the day before.

For convenience and following recommendations for DIG RNA labelling all the PCR products were gel purified with the band of the expected size cut out and purified using standard kit and

procedures. The purified product (2 µl) was run on a gel so that the product could be approximately quantified using comparison with the Eurogentec Smartladder (Figure 149). For a PCR product template with RNA polymerase promoter sequence it is recommended that 100-200 ng of PCR product be used.

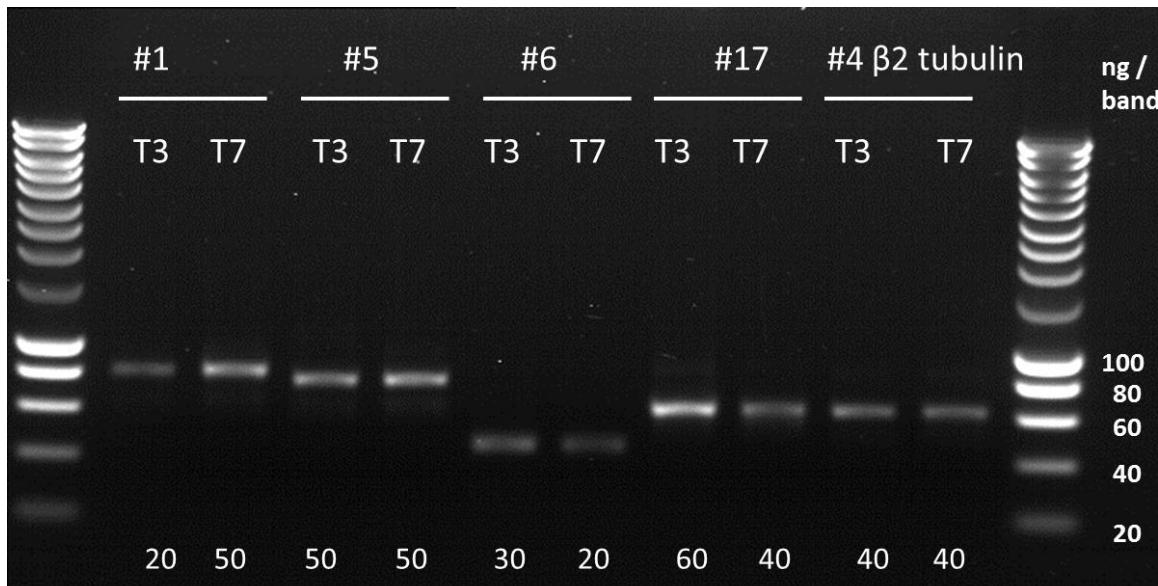


Figure 149. The Q5 PCR products shown in Figure 147 and Figure 148 were gel purified and quantified by comparison with Smartladder (the left ladder is half the default concentration, that shown by the right ladder). The diamondback moth testes specific and highly transcribed putative gene labels are given on the top and the type of RNA polymerase recognition site given with the prefix 'T'. The estimated quantity of DNA in 2 µl of PCR product is shown below each well.

8.2.25.5 DNA TO RNA

The DNA concentration of each PCR product was estimated and used to calculate how much PCR product should be used for efficient transcription of the DIG labelled RNA probe:

Table 46. The volume of PCR product, by reaction, that will provide 100-200 ng of DNA for the RNA probe synthesis.

| #1 | | #5 | | #6 | | #17 | | #4 | |
|-------|------|------|------|-------|-------|------|------|------|------|
| T3 | T7 | T3 | T7 | T3 | T7 | T3 | T7 | T3 | T7 |
| 14 µl | 6 µl | 6 µl | 6 µl | 10 µl | 14 µl | 6 µl | 8 µl | 8 µl | 8 µl |

The PCR products were added to:

2 µl of 10x concentrated DIG RNA Labelking Mix (Roche),

4 µl of 5x concentrated transcription buffer (Roche); and

2 µl of RNA polymerase (T3 or T7).

The reagents were mixed on ice and topped up to a final volume of 20 µl and incubated at 37 °C for two hrs in a thermocycler. The optional DNase step was avoided to minimise the risk of RNA

degradation by RNase contamination. The objective of this step is to make the RNA to DNA ratio as high as possible.

Following transcription those probes shorter than 500 nt in length were stored at -80 °C. Larger probes were subjected to hydrolysis as, according to the protocol, this produces stronger-specific staining and less non-specific staining.

8.2.25.6 PROBE HYDROLYSIS

The probes were hydrolysed at 60 °C:

Table 47. A description of the RNA probes length and hydrolysis requirements. The incubation time is dependent on the length of the starting RNA transcript and can be calculated by allowing ~ 15 min hydrolysis time for every 500 nucleotides of transcript.

| Probe # | #1 T3 | #1 T7 | #5 T3 | #5 T7 | #6 T3 | #6 T7 | #17 T3 | #17 T7 | #4 T3 | #4 T7 |
|--|------------------|------------------|------------------|------------------|------------------|------------------|-------------------|-------------------|------------------|------------------|
| Hydrolysis required (>500bp) | T | T | T | T | F | F | F | F | T | T |
| size (nt) | 734 | 734 | 706 | 706 | 466 | 466 | 472 | 472 | 574 | 574 |
| time (min) | 21 | 21 | 21 | 21 | - | - | - | - | 18 | 18 |

The hydrolysed RNA-DIG probes were precipitated using ethanol and re-suspended in 100 µl. A 5 µl sample of the hydrolysed probes were run on a 1.2% agarose gel with ethidium bromide at 120 V for 30 min. The non-hydrolysed probes were run on the same gel at 0.3 µl.

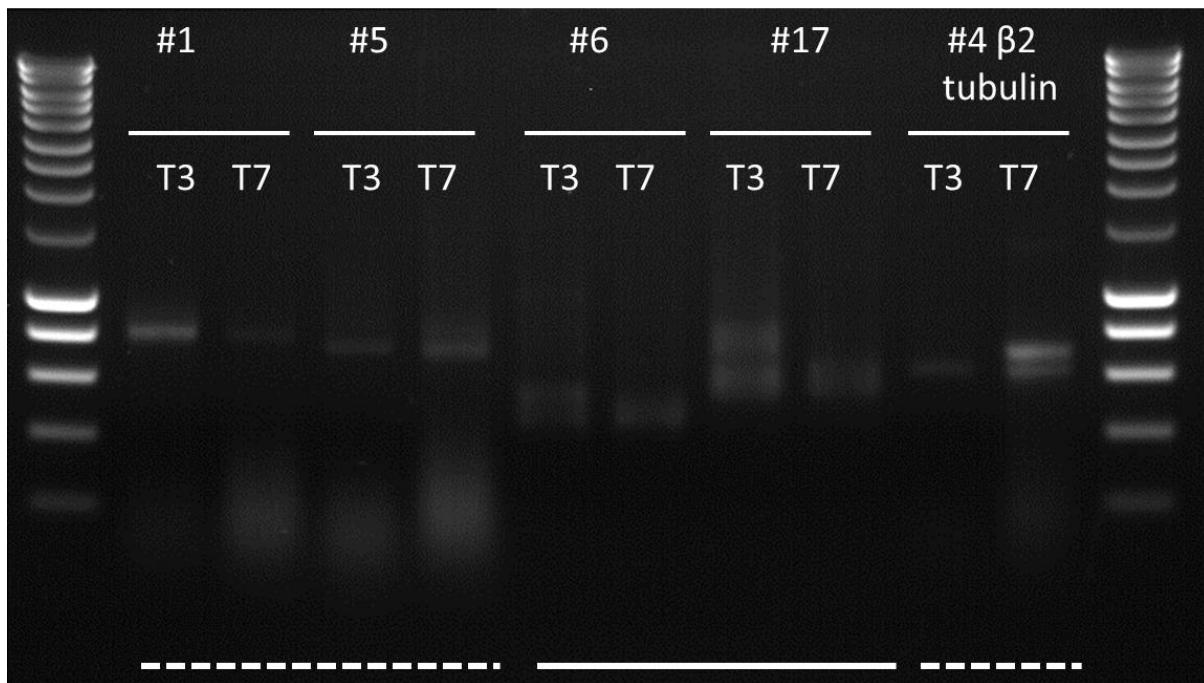


Figure 150. Those probes greater than 500 nt in length were hydrolysed (dashed white line underneath wells) those less than 500 nt were not (solid dashed lines). A 5 µl sample of the hydrolysed probes were run on a 1.2% agarose gel with ethidium bromide at 120 V for 30 min. The

non-hydrolysed probes were run on the same gel at 0.3 µl. Smartladder flanks the samples with 5 µl of volume. Two µl of x6 loading dye were used to load the RNA onto the gel. The gel was prepared using fresh reagents and equipment to avoid RNA-ase contamination.

The bands are unlikely to be DNA as the volume loaded is approximately one fifteenth of that loaded in Figure 149, ergo at this concentration the DNA would be undetectable. This suggests the bands are RNA and the smears near 200 bp in length are hydrolysed probe. The unusual double banding is a concern, however as we had no previous experience with this technique to know what effect RNA-DIG transcription had when visualised on a gel. However, as it must be RNA we were confident enough to proceed with ISH using these probes.

8.2.25.7 TESTES PREPARATION FOR *IN SITU* HYBRIDISATION

Following steps 21-30 in Morris *et al.*, (2009) testes from diamondback moth pupae were dissected, fixed (in the fix solution), digested (using proteinase K) and re-fixed as recommended. Testes were prepared a week before *in situ* hybridisation and were stored in hybridisation buffer at -20 °C prior to use.

For step 24 testes were digested for seven min with proteinase K due to a relatively lower surface area to volume ratio compared to *Drosophila melanogaster* testes.

8.2.25.8 ISH OF RNA PROBE TO DIAMONDBACK MOTH TESTES

The protocol was adhered to with changes at 35 and 42 where seven washes were applied rather than the default. For step 45, the addition of the staining solution, the duration in darkness was one hr for probes 9 and 10, one hr and a half for probes 5 and 6 and two hrs for probes 1, 2, 7 and 8 before the reaction was stopped.

8.2.25.9 MOUNTING THE SLIDES

The diamondback moth testes were more resilient than *Drosophila* but still need to be handled with care following steps 50-52 of the protocol. Some testes were exploded due to mishandling.

8.2.25.10 OBSERVATION

Slides were examined on an Olympus Bx50 upright microscope using DIC microscopy and x4 magnification. Every testes on each slide was examined, typically ten, with photos of the least damaged taken. Images were captured using a JVC KY-F75U three-colour CCD camera with KY-Link software. No image manipulation software was used. Higher magnification images were taken as appropriate and interesting.

8.2.25.11 TESTES DISSECTION

Two-day-old adult males were collected after being sexed at pupae. Testes were removed using fine tweezers and a light microscope. Testes were kept moist in testis buffer or PBS for viewing under fluorescent microscope to detect the appropriate marker. Water may be preferred as salt crystals do not form on evaporation. Testis status was confirmed by locating sperm bundles.

Fluorescence was assessed for the whole testis and following explosion by increased magnification inspection of sperm bundles.

8.2.26 MOLECULAR CONFIRMATION OF SEX FOLLOWING GDNA EXTRACTION

Accurate sexing of *Tribolium* will be critical in assessing efficacy of the designed construct mentioned above. Accordingly an experiment was designed to test the fidelity of a molecular method for discerning the sexes by the associated gDNA sequences. This would be also useful to test the sex of putative transgenics that die pre-pupae, to determine any sex-specific toxicity.

Sexing of *Tribolium* varies in reliability through the life stages with pupal sexing the most reliable due to presence of external genitalia. It is possible, but more difficult, to sex adult beetles and no method exists for sexing larvae by morphological differences between the sexes. A molecular method was therefore researched that relied on sex-specific genomic differences to reliably designate a sex to a given sample of gDNA, now possible following publication of the *Tribolium* genome.

Lagisz *et al.* (2010) provide a method which relies on a “set of three primers for multiplex PCR reaction” which was designed to result in “amplification of different length Y-specific and not-Y-specific (control) DNA fragments in a single PCR.” This was shown to work in six lab strains, successfully sexing at pre-pupae and adulthood. I conducted a blinded experiment to determine the sex of the gDNA of some test insects to prove the efficacy of the multiplex PCR (8.2.24).

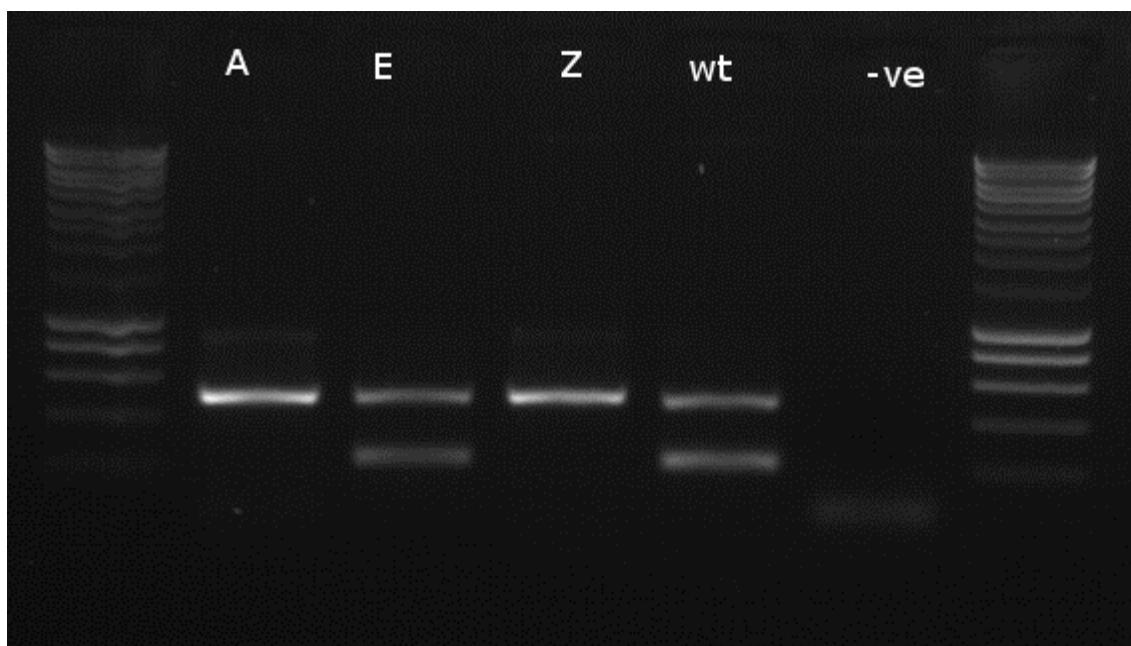


Figure 151. The multiplex PCR for sexing *Tribolium* gDNA works. Test of the sexing primers run on three transgenic lines (A, E and Z) and the wild-type (wt) GA-2 strain plus a negative control with no gDNA (-ve). A ~500 bp band was detected in males and females, a ~230 bp band was found exclusively in PCRs with male DNA as expected. Surrounding lanes provide DNA size standards. Sizes are as follows: from bottom, 200 bp increasing in 200 bp increments until 1000 bp, then 500 bp increments to 3000 bp (Eurogentec, Smartladder).

As shown in Figure 151, there were additional bands (~900 bp), noted by Lagisz *et al.*, (2010) but these can be ignored as our interest lies with the two informative bands. The primers

worked successfully and will improve identification of the sex of *Tribolium* at the level of the gDNA for those life stages that are difficult to sex by morphology. Throughout this thesis hundreds of *Tribolium* were successfully sexed using this method, as pupae and adults.

8.2.27 DIAMONDBACK MOTH 3' UTR ELUCIDATION USING BIOINFORMATICS

As an additional check, the gDNA region about the *β2-tubulin* unique axenome motif region extending into the putative 3' UTR and beyond (204 bp) were used in a Sequence Read Archive Nucleotide BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed: 26/03/2015). Our selected data were queried against DRX001985 RNA sequencing of *Bt* toxin-susceptible (PXS) diamondback moth strain (*Plutella xylostella* taxid:51655; study:DRP000685; submission:DRA000654; run:DRR002645). This yielded hits that extended into the putative 3' UTR. The read that extended deepest into the 3' UTR was aligned with our gDNA sequence and the 3' UTR elucidated (data not shown).

The Liverpool transcriptome resource was also inspected (Contig ID: >comp30984_c3_seq3 len=1485 path=[1:0-162 164:163-461 463:462-654 Length = 1485). In comparison to the previous alignment, the overall identity was lower, with less similarity towards the 5' and 3' ends. The available read extended only 26 bp into the 3' UTR.

8.3 GENERATION OF GERM-LINE TRANSGENIC STRAINS

8.3.1 PREPARATION OF TRANSFORMATION MEDIUM OR INJECTION BUFFER

Fifty ml LB medium was pipetted into a sterile tube. Filter sterilised 1M MgCl₂ and 1M MgSO₄ were added (625 µl). One ml of filter sterilised 20% glucose added and the total volume was homogenised and aliquoted into 15 ml centrifuge tubes. Stored at -20 °C or room temperature dependent on when required.

8.3.2 PREPARATION OF PLASMIDS FOR MICROINJECTION

Unless stated otherwise all injection mixes were comprised of 500 ng/µl of vector plasmid, 350 ng/µl of helper plasmid. The mixes were made up to 10 µl in total using 1 µl of injection buffer (5 mM KCL, 0.1 mM NaH₂PO₄, pH 6.8) and the appropriate volume of purified MilliQ H₂O to meet the concentrations given above.

8.3.3 PREPARATION OF INJECTION CAPILLARIES

Capillaries (Fisher, PMP-120) were made using a P-2000 laser based needle puller (Sutter Instruments, Novato, CA, USA) from aluminosilicate (with filament, catalogue number AF100-68-10, Intracel, Royston, UK). Capillaries were pulled ensuring a very fine tip with a steady taper – suitable programme details were developed by Adam Walker of Oxitec. The capillary were loaded with 3 µl of the injection mixture using an Eppendorf microloader. This volume was adjusted based on likelihood of the needle blocking or snapping so to avoid wastage of injection mix. The needle was tapped to remove air bubbles. The tip of the capillary was broken

open by rubbing it against the side of a glass slide (Fisher, FB58620) to ready for injection. The type of tip was adjusted depending on the task or species injected.

8.3.4 COLLECTION AND PREPARATION OF EMBRYOS FOR INJECTION

8.3.4.1 DIAMONDBACK MOTH

A 20mm glass slide was brushed with cabbage water on one side and allowed to dry. The slide was placed into a wild-type stock cage. It was left for between 5 to 30 min depending on the density of the eggs on the slide. Eggs were injected within two hrs of collection.

To ease the injection process, eggs could be washed off Parafilm coated with cabbage juice and realigned onto a fresh glass slide with a fine brush.

8.3.4.2 *TRIBOLIUM CASTANEUM*

Eggs need to be injected within six hrs of laying pre-cellularisation (Handel, 2000). A protocol has been developed by Berghammer *et al.* (2009). The information was supplemented with a personal communication from Gregor Bucher (2012) suggesting a humidity of above 40% and to limit the adult density relative to the mass of flour used – as too many insects can result in detrimental behaviour such as egg cannibalism. A review of Sokoloff's work supplemented this information.

Adult beetles were placed on pre-sifted white flour (max grain size 250 µm) for a two hr egg laying period at 30 °C. Adults were removed using a 710 µm sieve. Embryos were allowed to develop for an hr as they are sensitive to physical stress. The flour was then sieved using a 250 µm sieve. Eggs were immersed in room temperature tap water in order to remove the remaining flour. The eggs were gently swirled in the water. The remaining eggs were kept moist by immersion in shallow tray of tap water (eggs were still in sieve). A fine brush and microscope was used to move eggs from the sieve onto a glass slide, neatly aligned for ease of injection. Care was taken to avoid puncturing eggs with bristles. Post-injection the slides were placed into Petri dishes in a Tupperware box layered with 30% brine solution soaked tissue to maintain a high humidity for two to three days.

8.3.5 REARING INJECTED EMBRYOS

8.3.5.1 DIAMONDBACK MOTH

Slides with injected DBM embryos were moved to fresh Petri dishes two to three days post injection. Tissue paper was placed on top of the Petri dish and then the lid was secured preventing larvae escaping through the vents. Parafilm® was used to secure the lid. Upon hatching, artificial diet was added; the quantity depended on the number of larvae.

8.3.5.2 *TRIBOLIUM CASTANEUM*

After injection placed into a humid Tupperware box (>60 % relative humidity), tissue soaked in tap water layered underneath the glass slides. After two days at this humidity and 30 °C, eggs were placed into an equivalent but dry chamber. On emergence, in the next few days, larvae were placed into flour and yeast medium using a fine brush in small plastic pots.

8.3.6 CROSSING G₀ AND SCREENING G₁

8.3.6.1 DIAMONDBACK MOTH

Injections survivors (G₀) were collected as pupae, separated by sex and placed with wild-type pupae of the opposite sex at a 3:1 ratio. Females were pooled and males were used individually. Pupae were placed into plastic beakers with a piece of Parafilm coated in cabbage water (2 x 2 cm). Parafilm was replaced every two days and eggs were reared to pupae (G₁) to be screened using an Olympus SZX12 microscope with green, red and cyan filter sets (depending on the reporter protein). If survival were particularly high then pots could be set up with 30 G₀ and 10 wild-type females with three egg collections per pot.

8.3.6.1.2 OX5196 X OX4026

Crosses with wild-type were set up with one transgenic female to two males or one transgenic male to five wild-type females for both the putative OX5196 lines and the OX4026 stock population. This produced heterozygotes for each line. Heterozygotes were confirmed by screening for the main fluorescent marker for each construct (Table 48).

Table 48. The genotype and phenotype of our constructs to be crossed. Heterozygotes were produced by mating with wild-type. Insects fluoresce when using appropriate screening methods.

| Construct | Main marker | Phenotype | Genotype |
|-----------|-------------|--|----------|
| OX5196 | ZsGreen | Insects fluoresce green and express $\beta 2tubulin$ -tTAV in sperm. | Gg |
| OX4026 | DsRed2 | Insects fluoresce red and express AmCyan off tetracycline, regulated by TetO and tTAV. | Rr |

The crosses were expected to produce a 1:1:1:1 Mendelian ratio of GR:Gr:gR:gr (Table 49).

Table 49. A punnet square of the expected genotypes produced from the OX5196 and OX4026 cross. The letters represent the main markers and what fluorescent protein was screened for; G for ZsGreen and R for Dsred. Upper case the gene is present, lower case it is absent from the insect. We would expect only AmCyan positive sperm in GR. The G is synonymous for $\beta 2tubulin$ -tTAV and the R is for tetO-AmCyan expression.

| | |
|--------------------|----|
| OX5196 x OX4026 | Gr |
|--------------------|----|

| | | |
|----|----|----|
| gR | GR | gR |
| Gr | gr | |

8.3.6.2 *TRIBOLIUM CASTANEUM*

Either as young adults or pupae the injection survivors were sexed and then crossed one to one with another G₀ of the opposite sex. Crosses were conducted in small plastic pots with holes pricked into the lid. Two grams of diet were included. The incubator was maintained a temperature of around 32° C. At this rate the first eggs would emerge after 25 days or so (Howe, 1956). G₁ were screened for fluorescence using an Olympus SZX12 microscope.

The process of producing *piggyBac* mediated transgenic *Tribolium* varies from that in the literature (Berghammer *et al.*, 2009) so is summarised below in diagrammatic form (Figure 152). Use of egg baskets is unnecessary as the 250 µm sieve can be placed under a microscope and embryos carefully removed directly from the surface using a fine brush and placed onto a glass slide for injection. The flour and water form a paste facilitating injection.

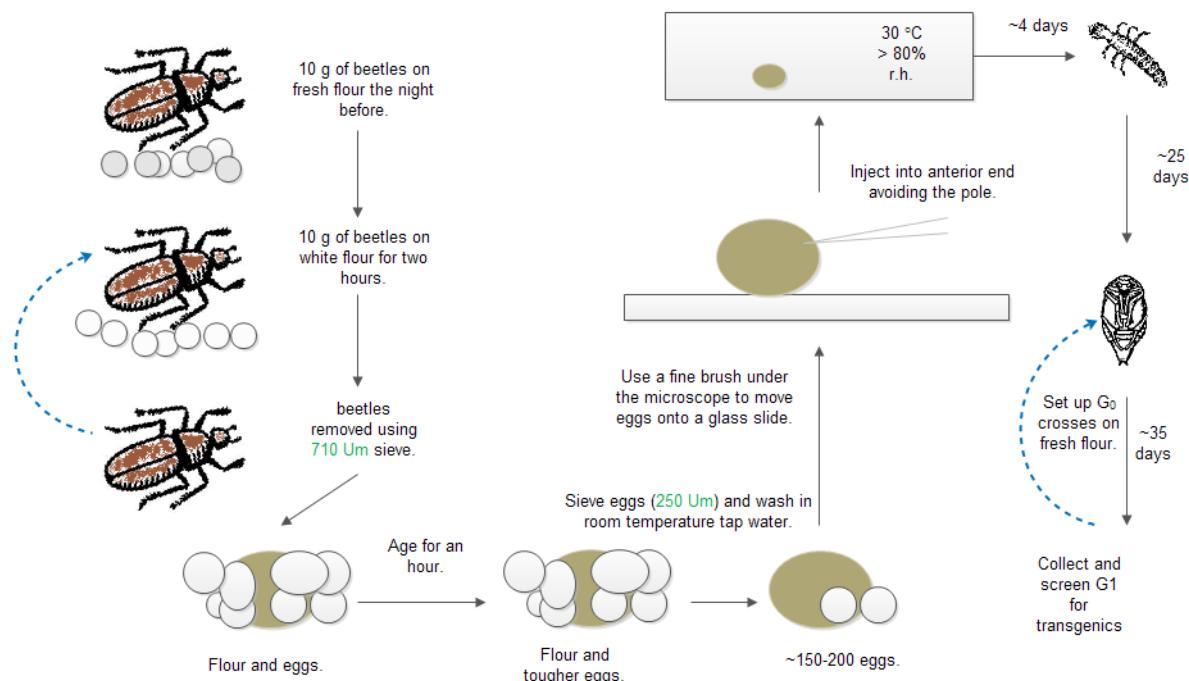


Figure 152. The method used to produce transgenic *Tribolium castaneum* differs from Berghammer *et al.* (2009). Time is saved by omitting the egg basket stage and removing the embryos straight from the sieve surface after washing. Each stage is described in the flow chart. Not to scale. Small grey circles represent flour and yeast diet, small white circles represent white flour, the brown oval represents the embryo (not true colours).

8.3.7 ESTABLISHING STOCK POPULATIONS

The transgenic positive pupae/larvae were selected and allowed to develop to adults. Upon maturation crosses with wild type insects were carried out to establish a stock population,

offspring were screened and transgenics separated. Long term stocks were based on a single insect (or sex crossed with wild type of the opposite sex) to reduce the likelihood that several independent insertions end up in one strain, these are referred to as lines or strains e.g. OX4698A and OX4698B. The transgenic insects were allowed to mate with the wild type in favourable mating ratios to ensure viable heterozygote offspring.

8.3.8 INJECTIONS

8.3.8.1 EMBRYOS

Microinjection of DNA is a straight forward yet laborious approach to generating transgenic organisms (Mehier-Humber & Guy, 2005). Embryos were injected into the posterior end using a microscope and micromanipulator with pneumatic pump (Figure 153). The targets are the nuclei that will go on to form the germ line, an area of the embryo known as syncytium (Daubenmire, 1936). Zygote viability correlates inversely with the size of the puncture caused by the capillary, determined by diameter and taper of the needle (Walton *et al.*, 1987).

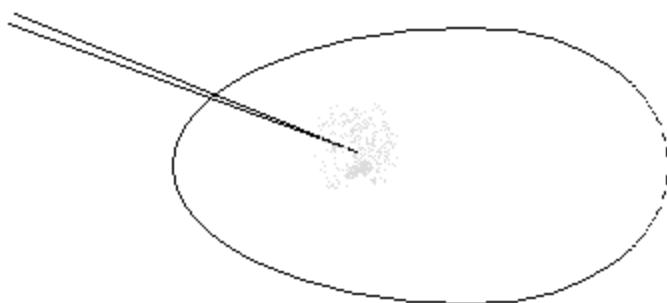


Figure 153. A diagram representing the "ideal" injection, the generic embryo is forced onto the needle at its posterior, air pressure is applied and the injection mix enters the embryo and diffuses through the cytoplasm at the approximate site of germline cell formation.

8.3.8.2 TRIBOLIUM CASTANEUM PUPAE FOR RNAI

Female pupae were collected and lined up on a glass slide with the tip of their abdomen attached by double side tape. Pupae were injected using a microscope and micromanipulator with pneumatic pump angled at 60°. The method is fully described in Posnien *et al.*, 2009. We used the same micromanipulator as for our embryo injections.

CHAPTER 9: APPENDICES

9.1 CONSTRUCTS

Constructs are displayed in a linear fashion to assist the reader. The constructs were injected as a circular plasmid with an additional source of transposase to facilitate integration.

Constructs were designed by Tarig Da'falla and / or myself. The synthesis method was designed by Tarig. Constructs were built by Caroline and / or Tarig. I was involved with the construction of OX4804.

9.1.1 CONSTRUCT DESIGN GENERAL COMPONENTS AND CONSIDERATIONS

9.1.1.1 PROMOTER

Transcription of eukaryotic genes that code for proteins is carried out by RNA polymerase II facilitated by general transcription factors. For spatial, temporal and tissue specific transcription, regulation is achieved by specific transcription factors (Levin & Tijan, 2003) facilitating transcription initiation complex formation. Accordingly endogenous promoter choice is crucial for desired specificity of transcription. The promoter is likely combinatorial in nature for genes with complicated expression patterns and timings. Transcription factor binding sites may be found both upstream and downstream of the promoter region respectively.

In some species the use of endogenous promoters to drive transgene expression are recommended (Lorenzen *et al.*, 2002) particularly if there is a large evolutionary distance between the construct-genetic-regulatory components and the species into which it is being inserted. This applies to *Tribolium castaneum* more-so than diamondback moth.

9.1.1.2 FIVE-PRIME UNTRANSLATED REGION

Use of endogenous promoters in isolation may not provide optimal transcription if key genetic components are missing. For example some genes have regulatory control in the 5' UTR or other non-coding regions of the gene (Barrett *et al.*, 2012). The 5' UTR is transcribed into mRNA but not translated into protein.

9.1.1.3 INTRONIC REGIONS

Introns are regions of DNA that are transcribed into pre-mRNA but are removed prior to mRNA formation by splicing (Barrett *et al.*, 2012). Introns have an important role in the regulation of gene expression with some genes not expressed at all if an intron is deleted (Rose & Beliakoff, 2000).

9.1.1.4 THREE-PRIME UNTRANSLATED REGION

Eukaryotic gene regulation occurs mainly at the transcription phase and can be enhanced or up-regulated using some well-known genetic sequences, such as simian virus 40 (SV40) (Banerji *et al.*, 1981), the 5' GTP cap and 3' polyA tail (the polyA tail is necessary for translation but may not increase stability; some stable mRNAs have very short tails (Clark & Pazdernik, 2012)). AU-rich elements play a critical role in the stability in some genes. These elements vary in size typically from 50-150 bp and contain multiple copies of the pentanucleotide AUUUA (Chen & Shyu, 1995).

9.1.2 OX4703

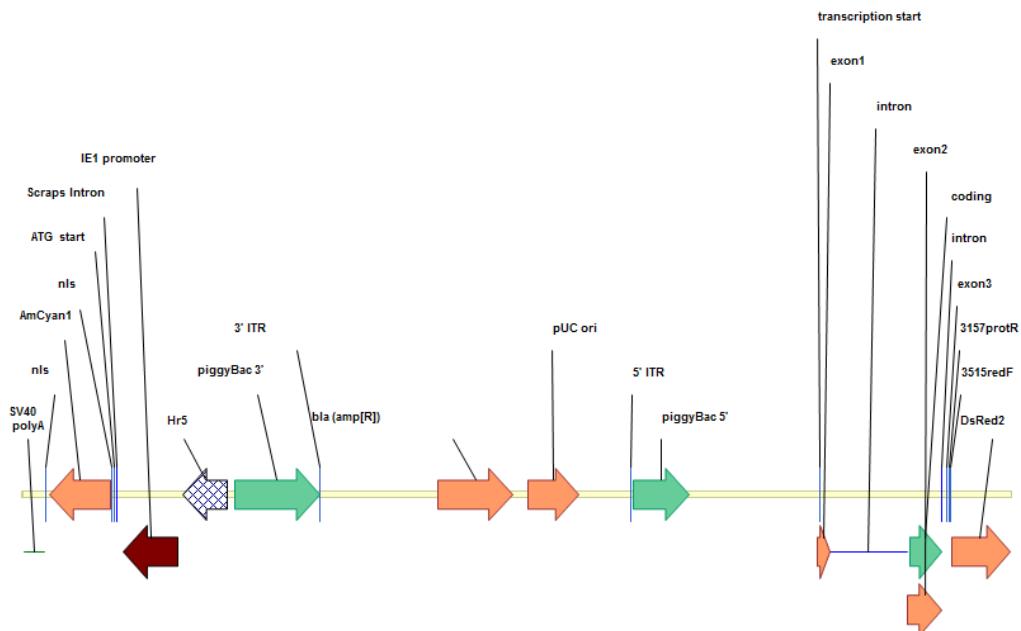


Figure 154. OX4703 PB-HrIE-Cyan-DBMb2tubpromoter-protamin-dsRed2-SV40 injected into DBM embryos and 11366 bp. The transgenesis module is on the left, the sperm marking module on the right. The exons 1-3 and introns therein are for the Dm protamine gene which is designed to be fused to the translated DsRed providing the fluorescent reporter protein with a DNA chaperone.

Table 50. Components listed in construct OX703 (After Oxitec internal document on OX4319).

| Name on plasmid map | Construct component | Component function | Donor | Detailed description |
|---------------------|---------------------|-------------------------|--|---|
| <i>piggyBac 5'</i> | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>Trichoplusia ni</i> (moth) | 5' end of <i>piggyBac</i> . <i>piggyBac</i> is a DNA (deoxyribonucleic acid) transposable element that, only when its ITR (inverted terminal repeats) are intact, is capable of integrating DNA flanking by |

| | | | | |
|----------------------|-------------------------------|---|---|--|
| | | | | element-specific DNA into other DNA through mediation of a transposase encoded by an ORF (open reading frame) within the element. Transformation was effected by introducing, with the transforming construct, a helper plasmid that supplied transposase activity but was itself unable to transpose into other DNA. This transposition-defective helper plasmid has an ORF encoding <i>piggyBac</i> transposase under the control of the <i>Drosophila melanogaster</i> hsp70 promoter. One of the inverted terminal repeats that flank the wild-type <i>piggyBac</i> transposase in <i>piggyBac</i> has been removed in the helper plasmid so that the helper plasmid cannot, itself integrate even though it encodes for active <i>piggyBac</i> transposase. |
| <i>piggybac 3'</i> | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>T. ni</i> (moth) | As above (Handler & Beeman, 2003) |
| SV40 polyA | Regulatory sequence 3'UTR | Stabilize mRNA | Virus | Regulatory sequence that helps stabilize mRNA. |
| nls | Nuclear localization sequence | Localises DsRed2 protein into the nuclei of cells | synthetic | NLS causes DsRed2 protein to accumulate within the nuclei of cells. Allows for spatial patterning of protein expression, which is useful to distinguish fluorescence patterns and distinguish from auto fluorescence. See Matins <i>et al.</i> , 2012. |
| DsRed2 | Protein coding sequence | Express DsRed2 | <i>Discosoma</i> Sp (coral) | The red fluorescent protein DsRed2 is a modified variant of DsRed. This allows the expression of a fluorescent protein. The transgenic insect with the marker gene fluoresces when excited by illumination of the appropriate wavelength. These fluorescent proteins, which have been used as markers in a wide range of vertebrate and invertebrate species, confer no known competitive advantage or disadvantage to the recipient, and no ecological or other consequences resulting from incorporation of these markers into the transgenic olive fly can be envisioned. |
| Scraps intron | Regulatory sequence for 5'UTR | Requirement for translation | <i>Drosophila melanogaster</i> | Stabilizes mRNA and required for translation of mRNA. |
| IE1/Hr5 | Enhancer /Promoter sequences | Control expression of DsRed2 | <i>Autographa californica</i> nuclear polyhedrosis virus (AcMNPV) | Promoter from <i>immediate-early-1</i> gene and <i>hr5</i> enhancer region. Hr5ie1, a fragment of the immediate-early-1 (ie1) gene with the Hr5 enhancer, from the <i>Autographica californica</i> nuclear polyhedrosis virus (AcMNPV). Hr5ie1 has been used in transformation markers in <i>Anopheles gambiae</i> (Grossman <i>et al.</i> , 2001), the Mexican fruit fly [<i>Anastrepha ludens</i> (Condon <i>et al.</i> , 2007)] and the Mediterranean fruit fly [<i>Ceratitis capitata</i> (Gong <i>et al.</i> , 2005)]. |
| Hsp70 | promoter sequence | Minimal promoter | <i>Drosophila melanogaster</i> | Minimal promoter to enable transcription of gene usually used adjacent to <i>tetO</i> in RIDL constructs. |

| | | | | |
|--|---------------------------------|---|--------------------------------|--|
| VP16 | Protein coding region | Component of tTAV protein | Herpes simplex (virus) | Component of synthetic transcription factor tTA |
| bla(amp^R) | Ampicillin resistance gene | Codes for the expression of β -lactamase. | | (Livermore, 1995). |
| Drosophila melanogaster protamine | Protein coding region | Codes for the expression of protamine | <i>Drosophila melanogaster</i> | Protamines replace histones during sperm development, enabling tightly packed DNA. The protamine recognizes and binds DNA. |
| pUC ori | Bacterial origin of replication | Plasmid replication point | <i>E. coli</i> | Allows high rate of replication of the construct within transformed bacterial cells. |

Diamondback moth $\beta 2$ -tubulin promoter was amplified using primers GF2750 (DBMtubAFL) and GF2751 (DBMtubKPN), using DreamTaq programme AMP-TD(55-52-50). This was purified and cut with *Afl*II and *Kpn*I. Vector #3867 was also cut with *Afl*II and *Kpn*I. The digests were gel extracted and ligated using rapid ligation kit with 5ul of vector and 4 ul of insert. Colonies were screened with GF#2750(DBMtubAFL) GF#2752 (RedR) and sequenced with primers TD#1431 (K10-short-1) and GF#2766.

9.1.3 OX4700

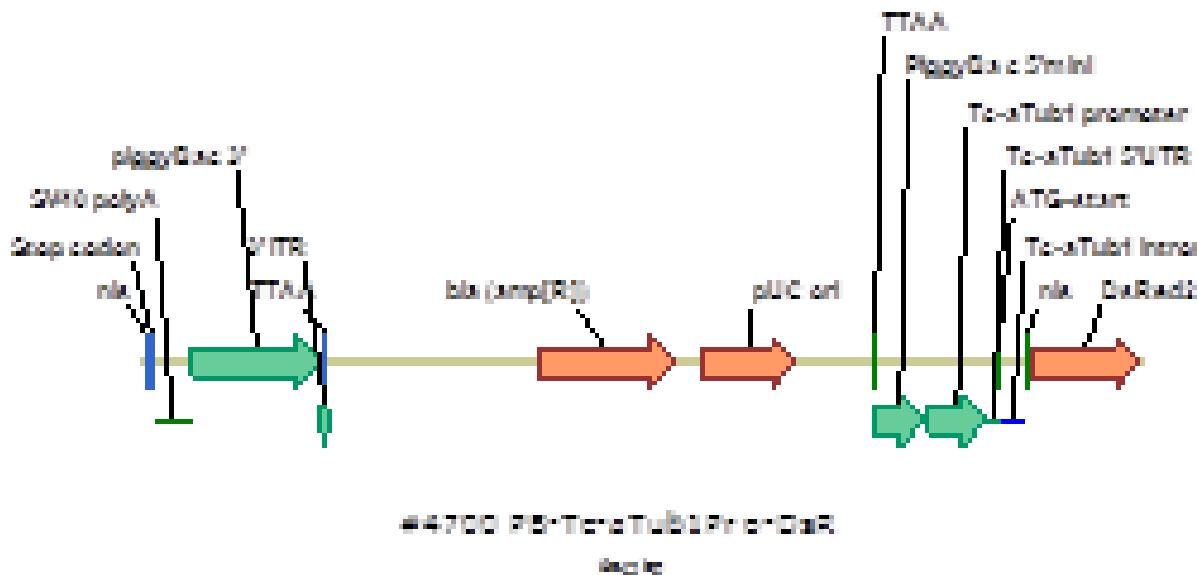


Figure 155. OX4700 PB-Tc α -tubulin-1-dsRed2-SV40 injected into Tc embryos at 6245 bp. This simple design was intended to confirm the α -tubulin-1 endogenous promoter driving DsRed expression as suitable for reporting transgenesis.

Table 51. Components listed in construct OX4700 (After Oxitec internal document on OX4319).

| Name on plasmid map | Construct component | Component function | Donor | Detailed description |
|---------------------|---------------------|--------------------|-------|----------------------|
| | | | | |

| | | | | |
|--|-------------------------------|---|--|--|
| <i>piggyBac 5'</i> | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>Trichoplusia ni</i> (moth) | 5' end of <i>piggyBac</i> . <i>piggyBac</i> is a DNA (deoxyribonucleic acid) transposable element that, only when its ITR (inverted terminal repeats) are intact, is capable of integrating DNA flanking by element-specific DNA into other DNA through mediation of a transposase encoded by an ORF (open reading frame) within the element. Transformation was effected by introducing, with the transforming construct, a helper plasmid that supplied transposase activity but was itself unable to transpose into other DNA. This transposition-defective helper plasmid has an ORF encoding <i>piggyBac</i> transposase under the control of the <i>Drosophila melanogaster</i> hsp70 promoter. One of the inverted terminal repeats that flank the wild-type <i>piggyBac</i> transposase in <i>piggyBac</i> has been removed in the helper plasmid so that the helper plasmid cannot, itself integrate even though it encodes for active <i>piggyBac</i> transposase. |
| <i>piggybac 3'</i> | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>T. ni</i> (moth) | As above (Handler & Beeman, 2003) |
| Tc α tubulin promoter | Promoter sequence | Promote transcription | <i>Tribolium castaneum</i> | Sequenced, described and tested in Siebert <i>et al.</i> , 2008. |
| nls | Nuclear localization sequence | Localises DsRed2 protein into the nuclei of cells | synthetic | NLS causes DsRed2 protein to accumulate within the nuclei of cells. Allows for spatial patterning of protein expression, which is useful to distinguish fluorescence patterns and distinguish from auto fluorescence. |
| DsRed2 | Protein coding sequence | Express DsRed2 | <i>Discosoma Sp</i> (coral) | This allows the expression of a fluorescent protein. The transgenic olive fly with the marker gene fluoresces when excited by illumination of the appropriate wavelength. These fluorescent proteins, which have been used as markers in a wide range of vertebrate and invertebrate species, confer no known competitive advantage or disadvantage to the recipient, and no ecological or other consequences resulting from incorporation of these markers into the transgenic olive fly can be envisioned. |
| bla(amp^R) | Ampicillin resistance gene | Codes for the expression of β -lactamase. | | Facilitates amplification in <i>E. coli</i> . |
| SV40 polyA | Regulatory sequence 3'UTR | Stabilize mRNA | Virus | Regulatory sequence that helps stabilize mRNA |

OX3905/6 was digested with NotI + BglII (4916 bp + 2084 bp). The product was purified on agarose gel and a normal spin column.

These following instructions were implemented:

1. Amplify TcATub1-Pro+5'UTR from *Tribolium* gDNA with 3022) TcTub-bsmb-f + 3023) TcTub-int-bsmb-r (= 639 bp). Digest with BsmBI and purify on agarose gel and a minElute column.
2. Amplify nls-DsRed2 from #3811 with 3024) Nls-bsmb-f3 + 3025) DsR-bsmb-r2 (= 752 bp). Digest with BsmBI and purify on agarose gel and a minElute column.

Ligate (3905-TcATub1-DsR). Screen bacterial colonies by PCR with 222) Diag-pb5 + 852) Diag4-DsRed (= 931 bp). Digest with BglII + KpnI (5174 bp + 1081 bp).

9.1.4 OX4804

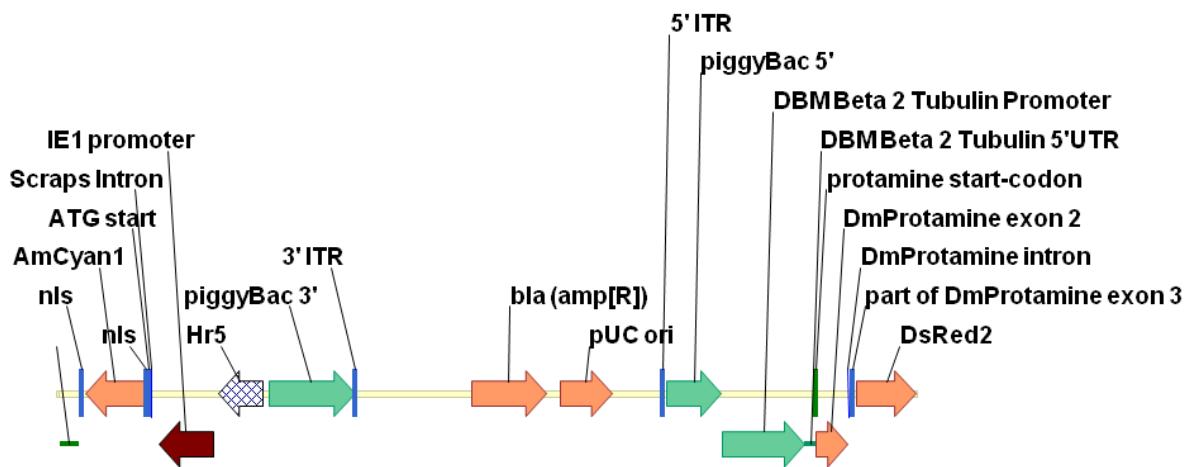


Figure 156. OX4804 PB-HrIE-Cyan-DBMb2tubpromoter-protamin-dsRed2-SV40 injected into DBM embryos, 9918 bp in length. Similar to OX4703 but with 5' UTR of the protamine removed. The transgenesis module is on the left, the sperm marking module on the right.

These following instructions were implemented:

Digest #4703 with XmaI + RsrII (7667 bp + 3651 bp). Purify on agarose gel and a normal spin column.

Amplify DBM-B2TubPro from pJET 1(3) with 3265) DBM-B2TubPro-bbs-f + 3271) DBM-B2Tub5utr-bbs-r2 (= 1108 bp). Digest with BbsI and purify on agarose gel and a minElute column.

Amplify DmProt-DsR from #4703 with 3272) DmProt-bsmb-f5+ 3268) DsR-fla-rsr-r (= 1224 bp). Digest with BsmBI + RsrII and purify on agarose gel and a minElute column.

Ligate (4703-B2Tub-Prot-DsR-corrected). Screen bacterial colonies by PCR with 3273) Diag2-DmProt + 2686) PB5-sexA-r (= 1382 bp). Digest with KpnI + SacI (7929 bp + 1576 bp + 413 bp).

9.1.5 OX4799

A positive feedback RIDL system where the design was intended to kill females off-tet.

To assist in tTA binding and transcriptional activation of the *tetO* promoter tTA has historically been fused to VP16 to create tTAV (Gallia & Khalili, 1998). VP16 is a 490-amino-acid protein that contains a core region and a C-terminal transcriptional activation domain. If the *tetO* is placed adjacent to a minimal promoter, for example the *hsp70* minimal promoter, then a transcription initiation complex will successfully form and drive expression of the downstream gene.

The *tetO* copy number (typically seven repeats) can also be increased to improve transcription efficiency (or reduced to retard).

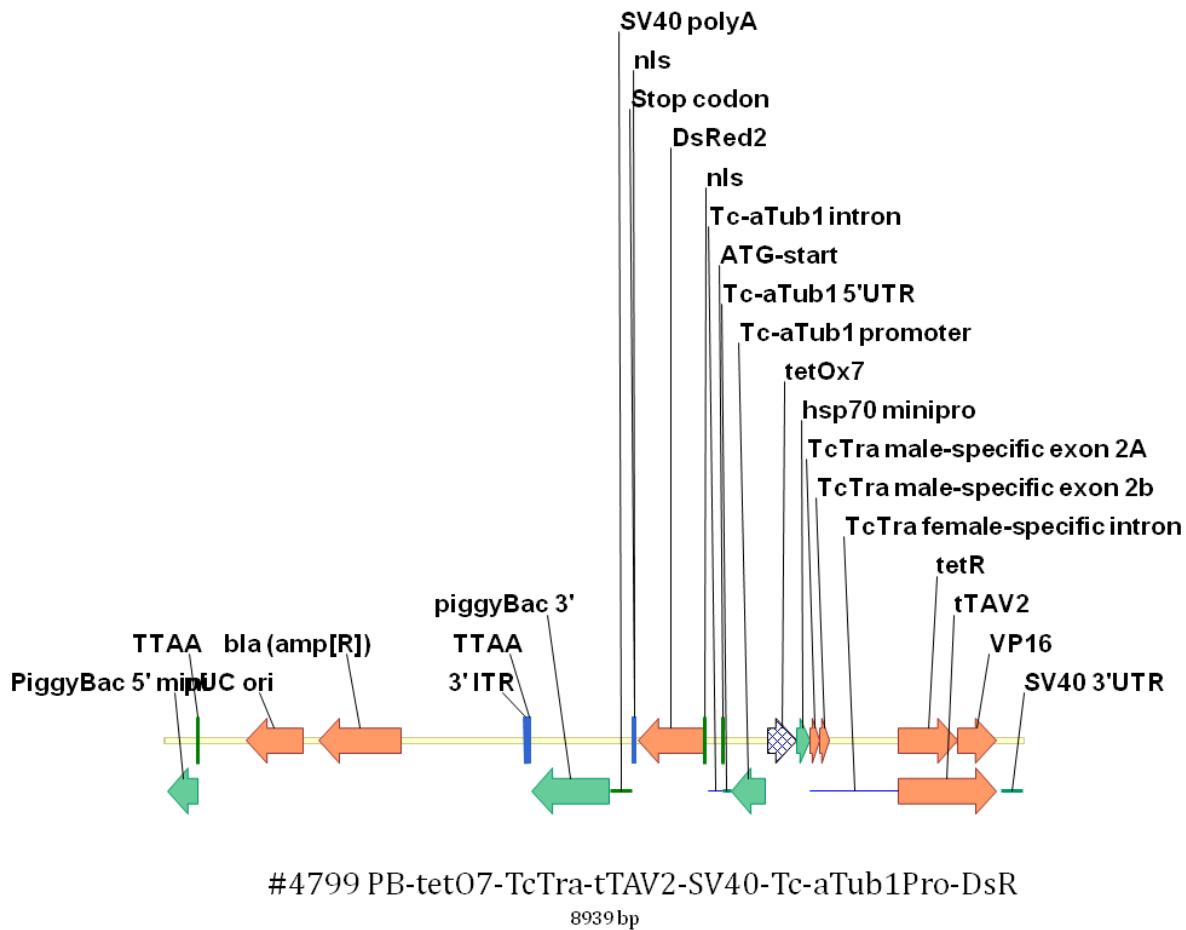


Figure 157. OX4700 a female lethal tetracycline regulated construct injected into Tc embryos at 8939 bp in length.

These following instructions were implemented:

Digest #4700 with NotI (linearize DNA). Dephosophorylate using Alkaline phosphatase and then heat inactivate.

Amplify tetO7-hsp70 minipro from #513 with 3248) teto7-bbs-f2 + 3249) hsp70-5utr-bbs-r (= 485 bp). Digest with BbsI and purify on agarose gel and a minElute column.

Amplify TcTra instron from Tribolium gDNA with 3250) Tc-int-bsa-f + 3251) Tc-int-bsa-r (= 949 bp). Digest with BsaI and purify on agarose gel and a minElute column.

Amplify tTAV2 from #4566 with 3252) tTAV2-bbs-f2 + 3253) SV40-bbs-r4 (= 1342 bp). Digest with BbsI and purify on agarose gel and a minElute column.

Ligate (4700-teto7-TcTra-tTAV). Screen bacterial colonies by PCR with 3254) TcaTubPro-f + 3255) TcTra-int-r (= 678 bp). Digest with NotI + KpnI (5995 bp + 2236 bp + 460 bp + 250 bp).

9.1.6 OX5067

Identical to OX4804 except for the replaced DBM *β2-tubulin* promoter region. The construct includes an extra 3613 bp in putative promoter region. The intention was to get a smaller fragment than this but due to a discrepancy between the strain which had been sequenced and our lab strain the addition of almost an extra 1 kb of sequence above that intended occurred.

Upon completion of designing the new construct electronically, the physical synthesis of the construct was initiated. The OX4804 putative promoter was removed using restriction enzymes (XmaI and Acc65I) on construct OX4804. The digested construct produced an 8908 bp (+ 1010 bp waste) sequence purified on agarose gel and a spin column.

To create the new extended promoter region, diamond back moth gDNA was amplified using the newly designed TD3443 and TD3444 (3372 bp) primers in a PCR. The product was digested with *BsaI* and purified on an agarose gel and minElute column. The band produced appeared larger than expected suggested our knowledge of the upstream region of the gene was incomplete.

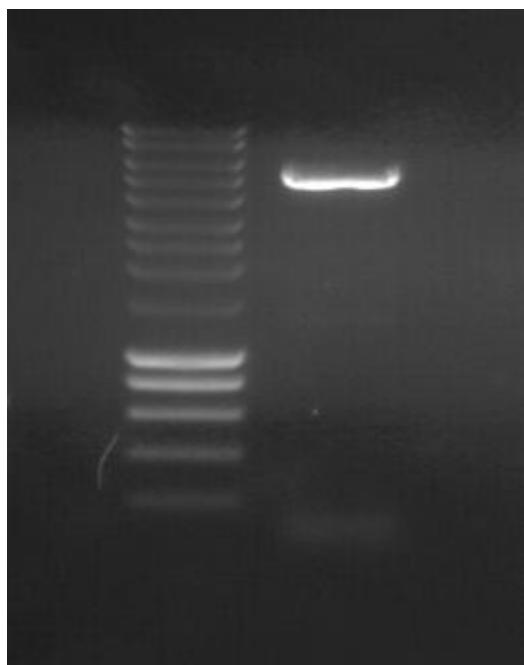


Figure 158. Gel electrophoresis at 1.2% agarose and 120V for 30 mins with eurogentec smart ladder. Amplification of the endogenous *β2-tubulin* promoter was achieved using TD3443 and TD3444 primers. The expected band size was 3372 bp, the band appears to be between 5000 and 6000 bp. However, it is distinctive and clearly defined suggesting successful amplification.

The extended *β2-tubulin* promoter construct was created by ligation of the two products using DNA ligase at 37 °C for 1 hr. The product of the ligation was used to transform *E. coli*. Forty-

eight colonies were picked and screened using a PCR reaction with primers specific to the new extended promoter region (TD3445 and TD3273). Sixteen of the forty-eight colonies were negative for expected amplicon (1424 bp). Six of the colonies were chosen for purification and miniprep formation. As an additional check, the miniprep product was digested with fast enzymes *KpnI* and *BamHI*. The digest product was run on a gel with expected bands of 9340 and 2914 bp in length. Two minipreps with strong digest bands were chosen for sequencing using five primers TD186, 3446, 3447, 3247 and 3273. Both sequencing runs had a 600 nt insertion when compared to our electronic version derived from the Liverpool diamondback moth genome data. This may be due to inter-individual genome variation and evolution between diamondback moth strains used.

9.1.7 OX5116

OX5116 represents a move toward simplicity in design, reducing the components in order to achieve testes specific expression. The protamine has been removed and the *β2-tubulin* gene component increased to include the 5' UTR, first exon, intron and part of the second exon (labelled collectively as DBM B2Tub intron; with the second exon a length of 18 bp) fused to the DsRed. The transgenesis module is on the left, the sperm marking module on the right.

The concept behind this construct was that although for *β2-tubulin* transcription to occur in many species studied required only a 100 bp or so upstream of the transcription site, these examples were all in the Diptera. There were some species that required some of the first exon or intron for enhancer activity and transcription to occur. The 5' UTR were included as well as the first exon, first intron and start of the second exon all fused directly to the DsRed. It was assumed if transcribed and translated the DsRed would be able to function as normal resulting in fluorescence in the sperm and testes.

These following instructions were implemented:

Digest #5067 with *KpnI* + *RsrII* (12294 bp + 1237 bp). Purify on agarose gel and a normal spin column.

1. Amplify part of B2Tubulin from DBM gDNA with 3658-B2Tub*KpnF* + 3659-B2Tub*BsmbR* (= 401 bp). Digest with *KpnI* + *BsmBI* and purify on agarose gel and a minElute column.
2. Amplify DsRed-SV40 from #5012 with 3660-DsRBsmb*F* + 3661-SV40*RsrR* (= 1004 bp). Digest with *BsmBI* + *RsrII* and purify on agarose gel and a minElute column.

Ligate (5067-B2Tub-DsR). Screen bacterial colonies by PCR with 3662-Diag1B2Tub + 825-Diag4-DsRed (= 586 bp). Digest with *BglII* + *KpnI* (11990 bp + 1051 bp + 610 bp).

Sequence with PR3662, PR94.

9.1.8 OX5118

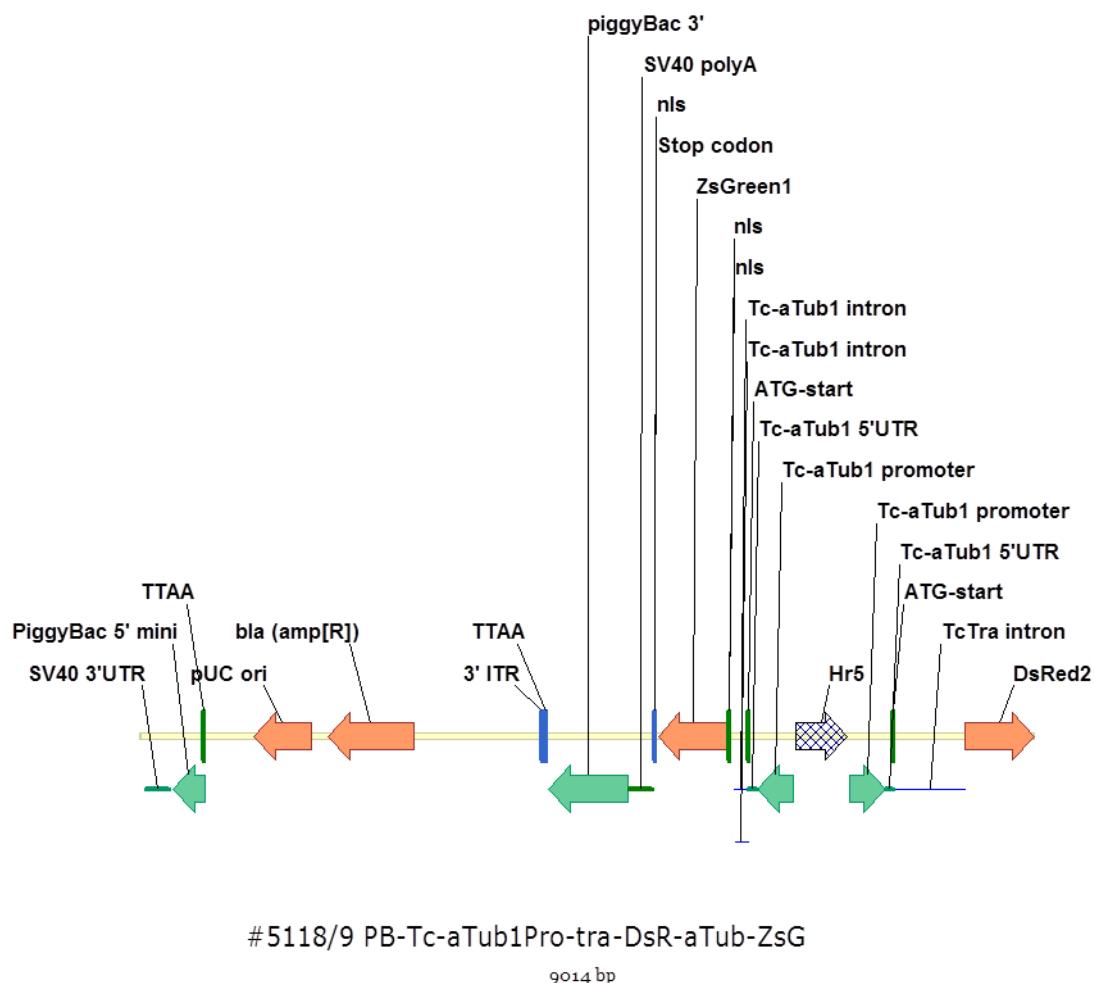


Figure 159. The construct OX5118 was designed to produce fluorescent green protein in both sexes and red females due to the splicing of the tra minigene. The Tctra intron was designed to be spliced from the female transcript producing DsRed. In males the intron would not be completely removed leaving in-frame stop codons preventing translation.

The reliability of DsRed2 as a transgenic marker in *Tribolium* was dubious and a replacement of ZsGreen sought. This was based on the difficulty in screening at all life stages and the lower than expected transformation efficiency. The re-design also simplified proceedings by removal of the feedback loop. The focus would be on testing on whether sex specific splicing of the transgene was occurring.

Table 52. Components listed in construct OX5118 (After Oxitec internal document on OX4319).

| Name on plasmid map | Construct component | Component function | Donor | Detailed description |
|---------------------|---------------------|-------------------------|--|---|
| <i>piggyBac</i> 5' | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>Trichoplusia ni</i> (moth) | 5' end of <i>piggyBac</i> . <i>piggyBac</i> is a DNA (deoxyribonucleic acid) transposable element that, only when its ITR (inverted terminal repeats) are |

| | | | | |
|--|-------------------------------|---|--|--|
| | | | | intact, is capable of integrating DNA flanking by element-specific DNA into other DNA through mediation of a transposase encoded by an ORF (open reading frame) within the element. Transformation was effected by introducing, with the transforming construct, a helper plasmid that supplied transposase activity but was itself unable to transpose into other DNA. This transposition-defective helper plasmid has an ORF encoding <i>piggyBac</i> transposase under the control of the <i>Drosophila melanogaster</i> hsp70 promoter. One of the inverted terminal repeats that flank the wild-type <i>piggyBac</i> transposase in <i>piggyBac</i> has been removed in the helper plasmid so that the helper plasmid cannot, itself integrate even though it encodes for active <i>piggyBac</i> transposase. |
| <i>piggybac 3'</i> | Vector sequence | Germline transformation | <i>piggyBac</i> from <i>T. ni</i> (moth) | As above (Handler & Beeman, 2003) |
| Tc α tubulin promoter | Promoter sequence | Promote transcription | <i>Tribolium castaneum</i> | Sequenced, described and tested in Siebert <i>et al.</i> , 2008. |
| nls | Nuclear localization sequence | Localises DsRed2 protein into the nuclei of cells | synthetic | NLS causes DsRed2 protein to accumulate within the nuclei of cells. Allows for spatial patterning of protein expression, which is useful to distinguish fluorescence patterns and distinguish from auto fluorescence. |
| DsRed2 | Protein coding sequence | Express DsRed2 | <i>Discosoma</i> Sp (coral) | This allows the expression of a fluorescent protein. The transgenic olive fly with the marker gene fluoresces when excited by illumination of the appropriate wavelength. These fluorescent proteins, which have been used as markers in a wide range of vertebrate and invertebrate species, confer no known competitive advantage or disadvantage to the recipient, and no ecological or other consequences resulting from incorporation of these markers into the transgenic olive fly can be envisioned. |
| bla(amp^R) | Ampicillin resistance gene | Codes for the expression of β -lactamase. | | (Livermore, 1995). |
| SV40 polyA | Regulatory sequence 3'UTR | Stabilize mRNA | Virus | Regulatory sequence that helps stabilize mRNA |
| ZsGreen | Protein coding sequence | Express ZsGreen. | <i>Aequoria Victoria</i> (jellyfish) | This allows the expression of a fluorescent protein under the control of a promoter. |

These following instructions were implemented by Caroline Phillips:

Digest #4799 with AgeI + BglII (8251 bp + 688 bp). Purify on agarose gel and a normal spin column.

Digest #3497 with AgeI + BglII (5047 bp + 715 bp). Purify on agarose gel and a minElute column.

Ligate (4799-ZsG). Screen bacterial colonies by PCR with 802) Diag-zsgreen + 3026) seq-TcATub1-5utr-1 (460 bp).

Digest 4799-ZsG with AscI + NheI (6581 bp + 2385 bp). Purify on agarose gel and a normal spin column.

1. Amplify Hr5 from #3497 with 3602-Hr5BbsF + 3603-Hr5BbsR (590 bp). Digest with BbsI and purify on agarose gel and a minElute column.
2. Amplify TcaTub1 from #4799 with 3604-TubProBbsF + 3605-Tub5utrBbsR (= 478 bp). Digest with BbsI and purify on agarose gel and a minElute column.
3. Amplify TcTra from #4799 with 3606-TcTraBsmbF + 3607-TcTraBsmbR (= 749 bp). Digest with BsmBI and purify on agarose gel and a minElute column.
4. Amplify DsRed from #513 with 3608-DsRBsmbF + 3609-DsRBsmbR (= 714 bp). Digest with BsmBI and purify on agarose gel and a minElute column.

Ligate (4799-ZsG-tra-DsR). Screen bacterial colonies by PCR with 3254) TcaTubPro-f (double the concentration) + 825) Diag4-DsRed (= 1890 bp + 667 bp). Digest with NheI + KpnI (6323 bp + 1606 bp + 1075 bp).

Sequence with PR3254, PR124, PR3700, PR38, PR3057.

9.1.9 OX5133

9.1.9.1 DESIGN

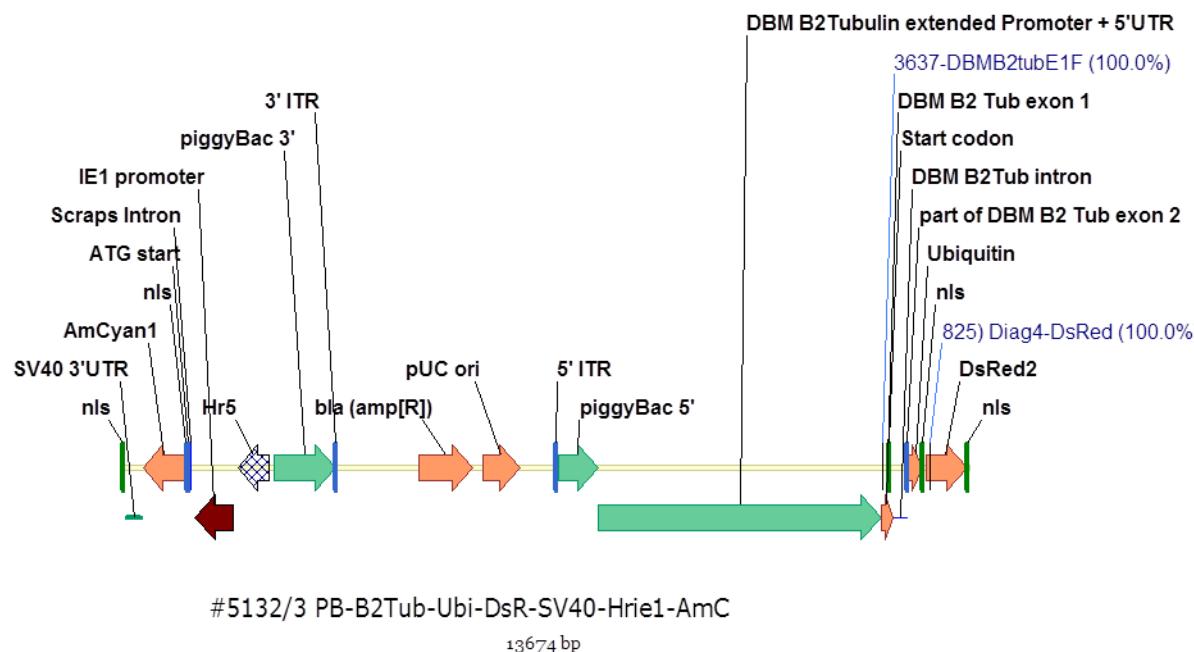


Figure 160. OX5133 was the first construct of the thesis to include ubiquitin, a component that cleaves the polypeptide during translation. This was intended to split the DsRed from the $\beta 2$ -tubulin ensuring proper folding of the protein. The transgenesis module is on the left, the sperm marking module on the right. The position of primers 3637 and 825 are also shown.

OX5116 showed testes specific transcription with the appropriate timing but no DsRed was detectable by microscopy. Accordingly it was assumed that the DsRed was getting interfered with by the $\beta 2$ -tubulin exons (24 amino acids) preventing correct folding and or function. A linker gene was needed to break these amino acid chains into discrete proteins so that DsRed could fold and tetramise normally. Ubiquitin was suggested as an appropriate linker molecule to serve this purpose.

9.1.9.2 OX5133 B2-TUBULIN AND DSRED SEPARATED BY UBIQUITIN

OX5116 was digested with *KpnI* and *BgIII* (12292 bp + 1055 bp). The products were purified on an agarose gel and a normal spin column.

These following instructions were implemented:

1. Beta2Tub CDS was amplified from OX5116 with primers TD3702-DBM-B2TubF and TD3703-B2TubBsmBR (= 428 bp). The product was digested with *KpnI* + *BsmBI* and purified on an agarose gel and a minElute column.
2. Ubiquitin was amplified from OX4656/7 (a hitherto unmentioned construct) with primers TD3704-UbiBsmB-F and TD3705-UbiBsmB-R (= 257 bp). The product was digested with *BsmBI* and purified on an agarose gel and a minElute column.
3. DsRed was amplified from OX4656/7 with primers TD3706-DsRtBsmBF and TD3707-DsR-BsmBR (= 799 bp). The restriction enzyme *BsmBI* was used to digest the product. The product was then purified on an agarose gel and a minElute column.

Following ligation (5116-Ubi-DsR) and screening of bacterial colonies by PCR with primers TD3637-DBMB2TubE1F and TD825) Diag4-DsRed (=783 bp), minipreps were digested with *ApaI* + *Kpn I* (12298 bp + 1376 bp) and *BglII* (12682 bp + 944 bp + 48 bp) for confirmation and sequenced with PR3057 and PR38.

9.1.10 OX5135

The same as OX5133 but with Dm protamine fused to the DsRed to assist in chaperoning sperm to DNA.

OX5116-Ubi- DsR was digested with *ApaI* and *AscII* (12909 bp + 765 bp). The digest was purified on agarose gel and a normal spin column.

The DmProt-DsRed was amplified from OX5067/8 with TD3708-DmProtBsmBF and TD3709- DsR-BsmBR (= 1193 bp). The product was digested with *BsmBI* and purified on an agarose gel and a minElute column.

Following ligation (5116-Ubi-Prot-DsR) and screening of bacterial colonies by PCR with TD975) Diag-ubiq and TD3273) Diag2-DmProt (= 332 bp). Minipreps were digested with *ApaI* and *BglIII* (12682 bp, 1390 bp and 2 bp) and sequenced for confirmation with PR975 and PR3057.

9.1.11 OX5145

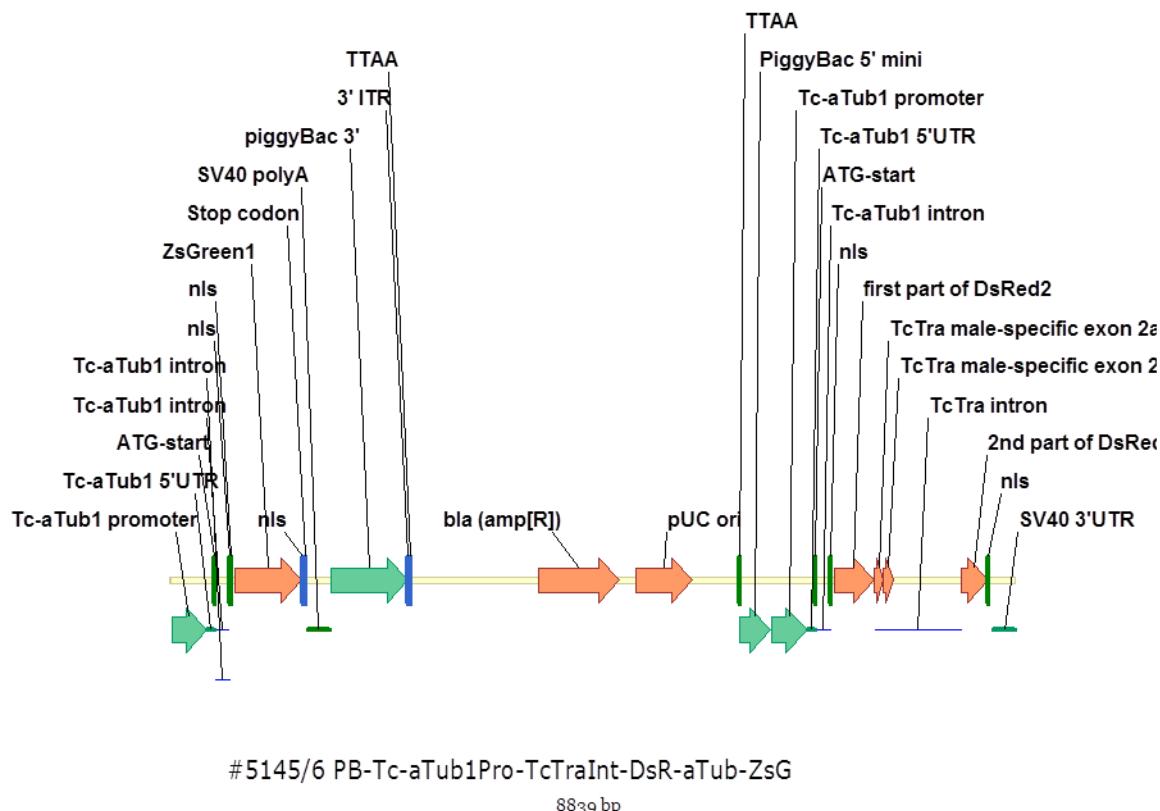


Figure 161. The construct OX5145 was designed to produce fluorescent green protein in both sexes and red females due to the splicing of the tra minigene. The Tctra intron was designed to be spliced from the female transcript producing functional DsRed. In males the intron would not be completely

removed leaving in-frame stop codons preventing translation and pushing the sequence out of frame.

These following instructions were implemented:

Digest #5118/9 with Ascl + NotI (6290 bp + 2724 bp). Purify on agarose gel and a normal spin column.

1. Amplify Tc-aTub1Pro-part of DsRed from #4700 with 3832)TcTub1BsmbF + 3833)DsRBsmbR (= 1090 bp). Digest with BsmBI and purify on agarose gel and a minElute column.
2. Amplify TcTra intron from Tribolium gDNA with 3834)TcTraBsmbF + 3835)TcTraBsmbR (= 945 bp). Digest with BsmBI and purify on agarose gel and a minElute column.
3. Amplify rest of dsRed2-SV40 from #4700 with 3836)DsRBsmbF + 3837)SV40BsmbR (= 596 bp). Digest with BsmBI and purify on agarose gel and a minElute column.

Ligate (5118-aTub1Pro-tra-int-DsRed). Screen bacterial colonies by PCR with 3810)DiagPB5long + 3255) TcTra-int-r (1374 bp). Digest minipreps with EagI (5661 bp + 1940 bp + 621 bp + 621 bp).

Sequence with PR3810, PR3057, PR2638, PR3838.

9.1.12 OX5196

This construct was designed to be crossed with OX4026. It uses ZsGreen as the main marker because OX4026 uses DsRed and AmCyan. This is the first of the constructs in diamondback moth and this thesis that drives tTAV expression using the $\beta 2$ -tubulin promoter and 5' UTR. The addition of the $\beta 2$ -tubulin 3' UTR to the 3' end of the tTAV is expected to provide translation in a sperm specific manner; something that has eluded us thus far. Expression of tTAV will be determined by crossing with OX4026 and reporter protein expression enhanced by *tetO*. This is the preferred protein detection method as protein work with fluorescent proteins proved inconclusive.

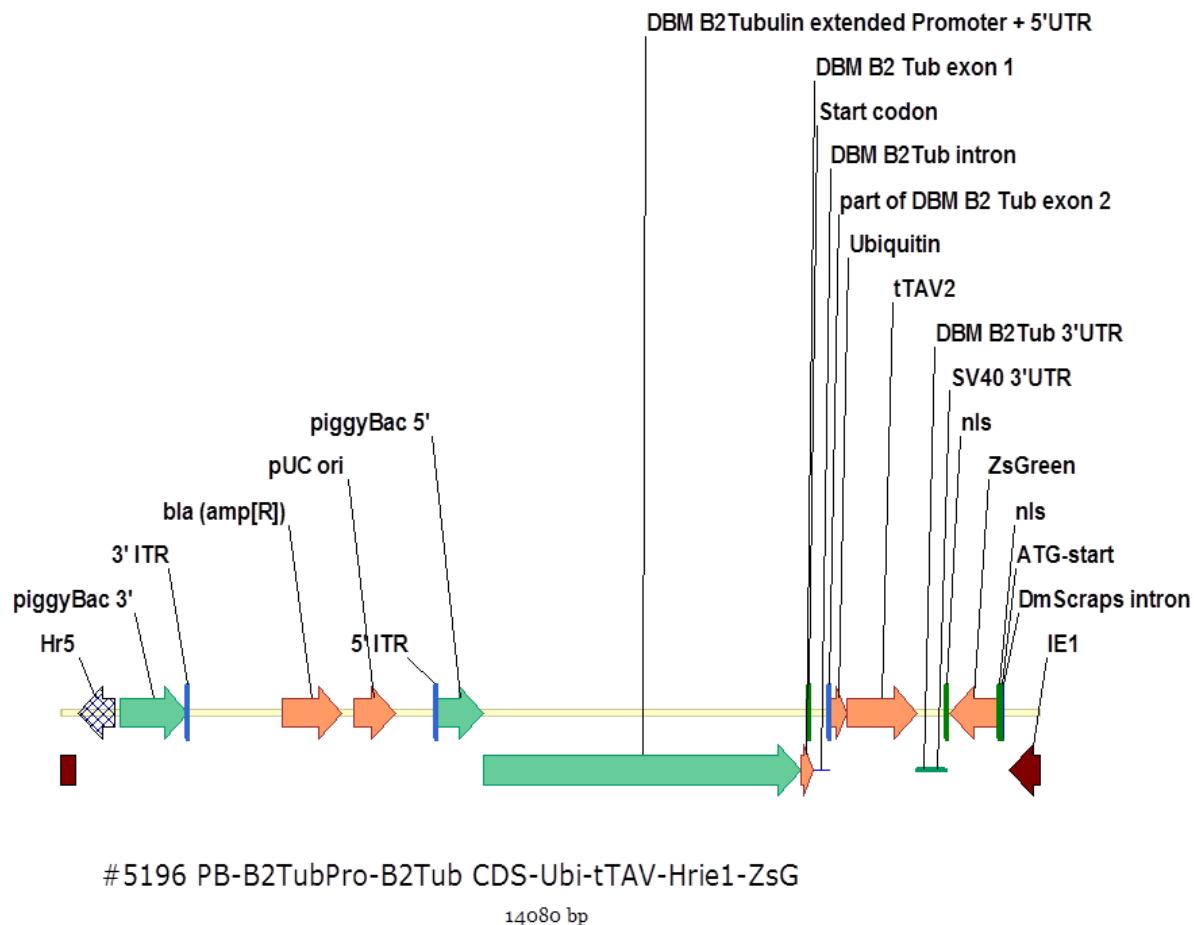


Figure 162. The construct OX5145 was designed to produce fluorescent green protein as a reporter protein. tTAV was expected to be expressed in the sperm. Transgenics would be crossed with OX4026 to report successful expression.

These following instructions were implemented:

Digest OX5133 with AvrII + NheI (= 11067 bp + ~~2607 bp~~). Purify on agarose gel and a normal spin column.

1. Amplify ubiquitin-tTAV from OX5034 with 3739-UbiBbsF + 4040)tTAVBbsR (=1269 bp). Digest with BbsI and purify on agarose gel and a minElute column.
2. Amplify DBM B2Tub2 3'UTR from pJET plasmid (from Mat) with 4041)B2TubBbsF + 4042)B2TubBbsR (= 196 bp). Digest with BbsI and purify on agarose gel and a minElute column.
3. Amplify part of IE1-ZsG-SV40 from OX3497 with 4043)IE1BbsF + 4044)SV40BbsR (= 1631 bp). Digest with BbsI and purify on agarose gel and a minElute column.

Ligate (5133-tTAV-ZsG). Screen bacterial colonies by PCR with 3268)DsR-fla-rsr-r + 4045)DiagB2Tub (= 1514 bp). Digest minipreps DNA with ApaI + NheI (11922 bp + 2157 bp).

Sequence with PR4045, PR2894, PR96, PR805.

9.1.13 OX5202

An attempt at a fsRIDL construct in *Tribolium castaneum*.

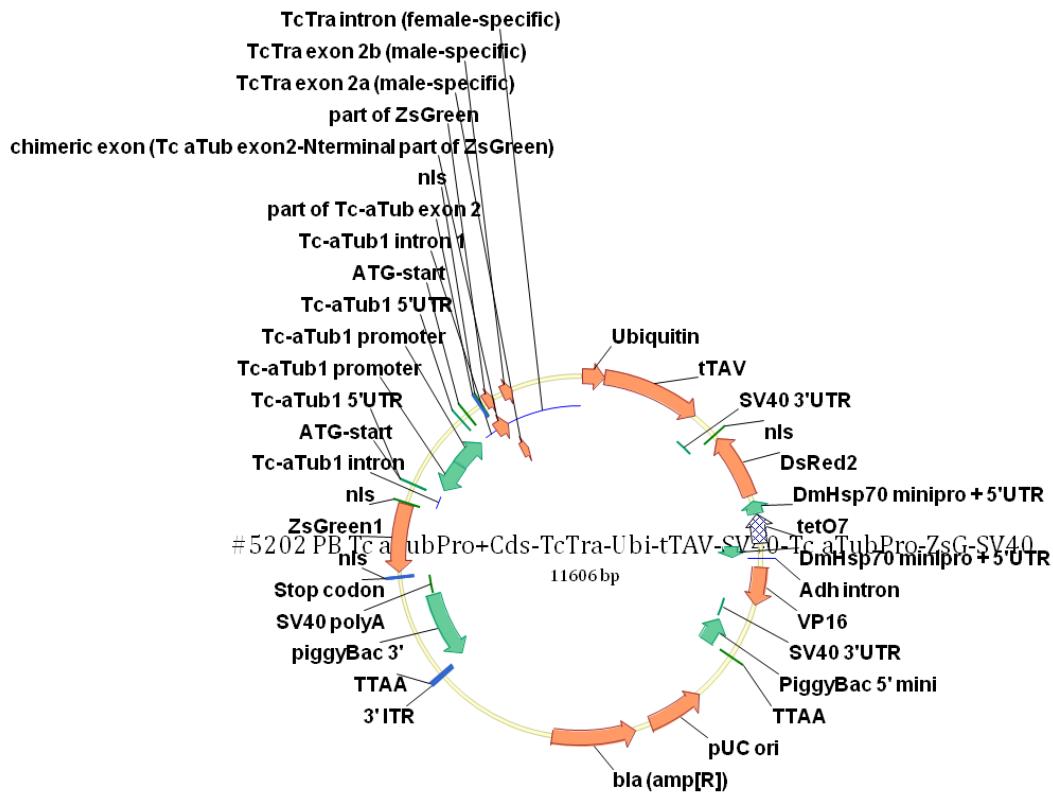


Figure 163. The construct OX5145 was designed to produce tTAV in females. In the absence of tetracycline this would bind to tetO and cause the expression of DsRed and VP16. Transgenics would be detected by screening for ZsGreen.

These following instructions were implemented by Caroline Phillips:

Digest #5118/9 with AsCI + NotI (6290 bp + 2724 bp). Purify on agarose gel and a normal spin column.

1. Amplify frag 2: Ubi-tTAV-SV40-part of DsRed2 from #5034 with 4004)UbiqBbsF + 4005)midDsRBbsR (= 1846 bp). Digest with BbsI and purify on agarose gel and a micro-column.
2. Frag 3:
 - a. Amplify part of DsRed2 from #5034/5 with 3998)DsRBsmBF + 3999)DsRBsmBR (= 515 bp). Digest with BsmBI and purify on a micro-column.
 - b. tetO7-hsp70 minipro from #5034/5 with 4000)hsp70BsmBR + 4001)teto7BsmBF (= 475 bp). Digest with BsmBI and purify on a micro-column.

Ligate (Frag 3 part of DsRed-tetO7-hsp70 minipro), amplify by PCR with 3998)DsRBsmBF + 4001)teto7BsmBF (964 bp). Digest with BsmBI and purify on agarose gel and a micro-column.

3. Amplify hsp70-VP16-SV40 (Frag 4) from #3952/3 with 4002)hsp70BbsF + 4003)SV40BbsR (= 904 bp). Digest with BbsI and purify on agarose gel and a minElute column.

Ligate (5118-ubi-tTAV-DsRed-teto-VP16). Screen bacterial colonies by PCR with 95) su-red-r + 3810)DiagPB5long (= 2086 bp). Digest with XhoI + NotI (6295 bp + 3636 bp).

Sequence with PR3254, PR2894, PR3057, PR4033, PR3810.

Digest 5118-ubi-tTAV-DsRed-teto-VP16 with XhoI + SpeI (9922 bp + 9 bp). Purify on a spin column.

1. Amplify Tc-aTubPro-part of ZsG from #5118/9 with 3994)aTuProBsmBF + 3995) ZsGBsmBF (= 786 bp). Digest with BsmBI and purify on a minElute column.
2. Amplify TcTra intron from #5145/6 with 3996)TctraBsmBF + 3997) TctraBsmBR (= 950 bp). Digest with BsmBI and purify on a micro-column.

Ligate (Frag1: Tc-aTubPro-part ZsG-TcTra int2) and PCR with 3994)aTuProBsmBF + 3997) TctraBsmBR (= 1566 bp). Digest with BsmBI and purify on agarose gel and a micro-column.

Ligate (5118-Tc-aTubPro-Tctra-tTAV-DsRed-teto-VP16-2). Screen bacterial colonies by PCR with 1031) Diag2-tTAV2 + 4007)TcTraIntF2 (= 321 bp). Digest with XhoI + SpeI (9922 bp + 1684 bp).

Sequenced with PR1102, PR3255.

This construct was difficult to synthesis and took almost six months to build, slowing the *Tribolium* research programme.

9.1.14 OX4026

This construct was a reporter line already in stock maintained by Debs Granville. Initial transformation was achieved by Adam Walker.

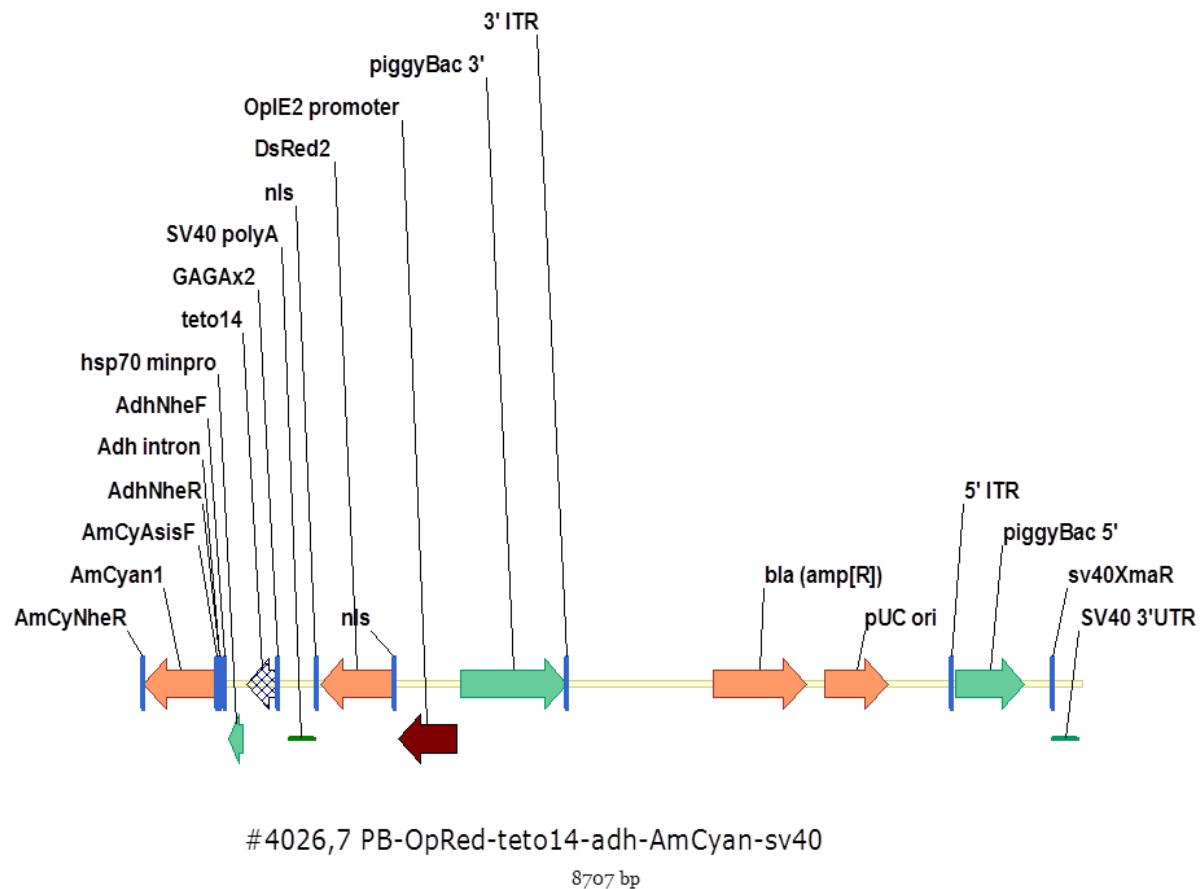


Figure 164. The construct OX5145 was designed to produce fluorescent green protein as a reporter protein. tTAV was expected to be expressed in the sperm. Transgenics would be crossed with this construct OX4026, to report successful expression. OX4026 has DsRed as the transformation marker. tTAV will bind tetO and enhance AmCyan expression with the assistance of the minimal promoter hsp70.

9.2 5' RACE OF 5'UTR DIAMONDBACK MOTH B2 TUBULIN

Male testes RNA were extracted from wild-type diamondback moth. The quality of the male testes RNA was assessed using gel electrophoresis and quantified using a spectrophotometer. No DNA contamination was apparent from the absorbance by wavelength plot (one peak at 260 nm).

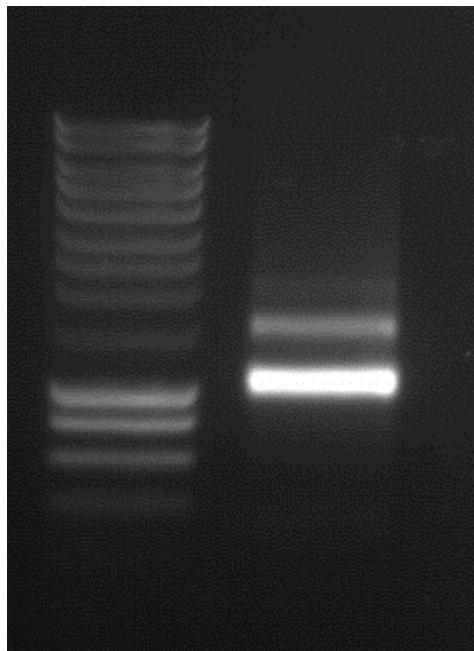


Figure 165. RNA integrity assessed prior to cDNA synthesis. Both the 18S and 28S rRNA appeared as sharp bands after electrophoresis of total RNA in 1.2% agarose gel at 120 V for 30 minutes. The 28S is approximately twice the intensity of the 18S rRNA. Smearing of RNA is negligible. The DNA smart ladder is included alongside the RNA sample.

The mRNA was converted to first strand cDNA using the SMARTer RACE cDNA amplification kit. After first and second strand synthesis the ds cDNA was purified, cloned and sent for sequencing. The sequence was compared to the transcriptome database using Vector NTI and found to be 19 nt and 24 nt shorter (two bacterial colonies that were positive for the sequence to be amplified from the purification and amplification step with no remarkable differences. An 8 nt sequence not found in the transcriptome was also present. This was assumed to be a relic of the adaptor sequence used in the 5' RACE.

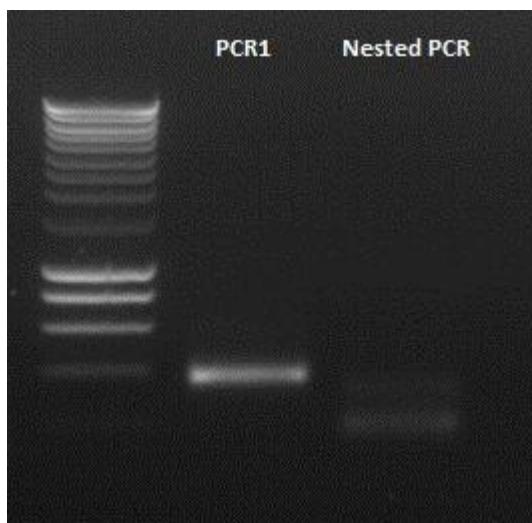


Figure 166. The cDNA product of 5' RACE of testes derived mRNA from diamondback moth. The primary PCR reaction was successful producing a distinct band approximately 400 bp in size. The nested PCR failed probably due to incorrect annealing temperature. For PCR1 the gene specific primers TD3450 and TD 3451 were used. Touchdown PCR (adjustment of annealing temperature)

was used to improve specificity in the earlier cycles and amplification in the latter cycles. The clean band was deemed suitable for purification, amplification and sequencing.

9.3 PRIMER SEQUENCES

Primers in this thesis were prefixed by the researcher's initial placing the order for the primer to be synthesised (Table 53).

Table 53. The full primer list for this thesis providing the primer name, nucleotide sequence and author. The case of the sequence represents it's sequence. For example SS100 for Sarah Scaife's 100th primer ordered. If not initials are provided then default to assume it is TD. The name gives a description as to the location of the primer relative to a gene of interest. Often the species is provided as DBM for diamondback moth or Tc for *Tribolium castaneum*.

| Primer number / name | Sequence | Author |
|--------------------------------|---|---------------|
| 2753 DBM Tub2 Rev | CTCTTCCTCATCAAACCTGCCCTC | Fu, G. |
| 2841 Beta 2 for across intron | ATTGGAGCTAAGTTCTCGGGAAGTGATATC | Fu, G. |
| 2842 Protamine intron F | CGCATCTCAATGAATTACAGAACTG | Fu, G. |
| 2843 Protamine intron R | CCATCCGACATCGAGACAAACAC | Fu, G. |
| 2845 Protamine ex2.ex3 forward | CGCCGGATGGCATGCAA | Fu, G. |
| 100 PB 5'end | CCACGAGGCGTAGCCGAG | Scaife, S. |
| 164 REDSEQR | GGTGATGACGTTCTCGGAGG | Scaife, S. |
| 306 DMMST35BRTF | GAGTGCAAGAGCCTGTGG | Scaife, S. |
| 376 Dsred-DraF | GGTGTGCACCTGGTGCTGCGCTGCGCGGTGGT ATGGCCTCCTCCGAGAACGTC | Scaife, S. |
| 476 TetNotF | CGGCGCTCGCGGCCGC ATCGATCTCTATCACTGATAAGGGAGG | Scaife, S. |
| 857 TaqVp16F2 | CCACGCCGATGCCCTGGA | Scaife, S. |
| 858 TaqVp16R2 | GGTGAACATCTGCTCGAACTCGAAATC | Scaife, S. |
| 1487 dbm 17StaqF | ctaggatggcccacaccggtgatac | Scaife, S. |
| 1488 dbm 17StaqR | ccgggtatcaccggtgtggcccatc | Scaife, S. |
| 8 Diag-DsRed2 | CTGGGAGGCCTCCACCGAGC | Scaife, S. |
| 38 DIAG2-DSRED | CTTCAGCTTCACGGTGTGTCGGC | Scaife, S. |
| 95 su-red-r | CTTGGCCATGTAGATGGACTTGAACCTCC | D'afallaa, T. |
| 118 su-cyan-r1 | CTTGAAGGGCGACGTGACCGCC | D'afallaa, T. |
| 193 Diag2-pb5 | GCGCGAATCCGTCGCTGTG | D'afallaa, T. |

| | | | |
|------|------------------|---|---------------|
| 222 | Diag-pb5 | CTGATTTGAAC TATAAC GACCGCGTG | D'afallaa, T. |
| 345 | 3'cent-3 | AAACCTCCCACACCTCCC | D'afallaa, T. |
| 802 | Diag-zsgreen | GTACTCGGTGAACACGC GG | D'afallaa, T. |
| 825 | Diag4-DsRed | CTCGATCTCGAAC TCGTGGC | D'afallaa, T. |
| 975 | DIAG-UBIQ | CGAGCGATA ACCATCGAGAAC | D'afallaa, T. |
| 1032 | Seq3-tTAV2 | CGTTGGAGCT GTTGAACGAAG | D'afallaa, T. |
| 1128 | DIAG2-SV40 | GATGAGTTGGACAAACCACA ACTAGA | D'afallaa, T. |
| 1405 | Diag7-DsRed7 | GCCACCGAGTT CGAGATCGAG | D'afallaa, T. |
| 1406 | Diag7-DsRed2 | CCATGGTCTT CTTCTGCATCAC | D'afallaa, T. |
| 1431 | K10-short-1 | ttgcgattaccagt gattt g | D'afallaa, T. |
| 1724 | TTAV-BSMB-R | GAGATGGACTTTGGCACCGTC | D'afallaa, T. |
| 1830 | BM-3UTR-R2 | AGCCAAATCAGTT CAGCCGTTCTTGTG | D'afallaa, T. |
| 2797 | DIAG-QS2 | ctgatcgctcgtaagt gctg | D'afallaa, T. |
| 3000 | Tc-dsx-f1 | agtctggactgcgactcctc ag | D'afallaa, T. |
| 3001 | Tc-dsx-f2 | ctcctcgcagt gctccaa c | D'afallaa, T. |
| 3002 | Tc-dsx-r1 | CCCTCGCATCCTTCAGGATG | D'afallaa, T. |
| 3003 | Tc-dsx-r2 | CGGGAAAGCCTCTTCCAAGTC | D'afallaa, T. |
| 3004 | Tc-dsx-r3 | GGCATATTGAGAATAACGTCCACTAG | D'afallaa, T. |
| 3005 | Tc-dsx-r4 | CTGCCTCCACGCCCTCGTC | D'afallaa, T. |
| 3006 | Tc-dsx-r5 | GCAGGATGATACATGGACATGGA | D'afallaa, T. |
| 3007 | Tc-dsx-r6 | TGTGGGGACTGATCCTAACAGAG | D'afallaa, T. |
| 3022 | TcTub-bsmb-f | agcgcgtctcaggccgactgc agtgaacggttatgatg | D'afallaa, T. |
| 3023 | TcTub-int-bsmb-r | AGCGCGTCTCTGTGGATT CACGCTGGAAAGGAAAACG | D'afallaa, T. |
| 3024 | Nls-bsmb-f3 | agcgcgtctcaccaccatggatcccacccacccaaga | D'afallaa, T. |
| 3025 | DsR-bsmb-r2 | AGCGCGTCTCAGATCTCAGGAACAGGTGGTGG | D'afallaa, T. |
| 3026 | seq-TcATub-5utr | gtgggttaattgtaaaactcaactacc | D'afallaa, T. |
| 3057 | Diag8-dsred | aacgtcatcaccgagg tcatg | D'afallaa, T. |
| 3184 | TcTra-e1-f1 | atgtcgggttcaaaggagcca | D'afallaa, T. |
| 3185 | TcTra-e1-f2 | ggaacacccgaa gttcaatgca | D'afallaa, T. |
| 3186 | TcTra-e3-r1 | ACTCCTGGCTTGT CGTTCC | D'afallaa, T. |
| 3187 | TcTra-e3-r2 | GCGATGTCTGCCGGACT | D'afallaa, T. |
| 3248 | teto7-bbs-f2 | acgcgaagacatggccgaggcgc gcccagg ttcg | D'afallaa, T. |

| | | |
|------------------------------------|---|---------------|
| 3249 hsp70-5utr-bbs-r | TCGCGAAGACATGTGGTACCTGCAGATTGTTAGCTTGT | D'afallaa, T. |
| 3250 Tc-int-bsa-f | acgcggctcaccaccatggttaattggtttcatagtcttgacg | D'afallaa, T. |
| 3251 Tc-int-bsa-r | ACGCGGTCTCACCATCTGAAATCAAGCAGTTTCTCAT | D'afallaa, T. |
| 3254 Tc α tubulin promoter | cagctgaggtcccggtcgt | D'afallaa, T. |
| 3255 Tc transformer intron reverse | TCCGGTCGCGATGCTGCATT | D'afallaa, T. |
| 3265 DBM-B2TubPro-bbs-f | acgcgaagacatccggaaatggtagttatacgattaaaggctatgg | D'afallaa, T. |
| 3266 DBM-B2Tub5utr-bbs-r | TCGCGAAGACTCAGAGCATTTCATTATAATAAACACTAGA GTCTCC | D'afallaa, T. |
| 3267 DmProt-bsmb-f4 | agcgcgtctcaactttttaaaattagttgtgacaa | D'afallaa, T. |
| 3268 DsR-fla-rsr-r | TGTCCAAACTCATCAATGTATCTAACGCGAGT | D'afallaa, T. |
| 3269 Diag-DmProt | TCGTTTATTCACTTACCCATTGAAG | D'afallaa, T. |
| 3270 DmProt-int-seq | AATTACACGTCTGATTGATCAAC | D'afallaa, T. |
| 3271 DBM-B2Tub5utr-bbs-r2 | TCGCGAAGACTCTCATTTCAATTATAAACACTAGAGTC TCC | D'afallaa, T. |
| 3272 DmProt-bsmb-f5 | agcgcgtctaatggatcaaataatgtaaatgagtgc | D'afallaa, T. |
| 3273 Diag2-DmProt | TGCACAGGACTTTATTGGCTTACA | D'afallaa, T. |
| 3275 Diag4-DmProt | gaaagaagcactgtgacttgaagc | D'afallaa, T. |
| 3372 Tc-sexing-F0 | TTTGTACAGCAAAGGACGCCGA | D'afallaa, T. |
| 3373 Tc-sexing-R1 | CATCTGCACTCAAAGCGCTGTCAA | D'afallaa, T. |
| 3374 Tc-sexing-R2 | GGCAAAAAATTACCCGTCTGGAG | D'afallaa, T. |
| 3443 DBM-B2TPro-bsa-f | acgcggctcaccggagatcgagcctcatcagaccaag | D'afallaa, T. |
| 3444 DBM-B2T-5utr-bsa-r | ACGCGGTCTCGGTACCTCCCTGATCAGCTCACT | D'afallaa, T. |
| 3445 Diag-DBM-B2T | gtgtggataatggctacctctcg | D'afallaa, T. |
| 3446 DBM-B2T-seq-1 | TGTAACATACGTGTAGGTACGCA | D'afallaa, T. |
| 3447 DBM-B2T- | CTAATAACGTAGGGAGGTTAGATCA | D'afallaa, T. |

| | | |
|---------------------|---|---------------|
| seq-2 | | |
| 3450 DBM-B2-5RACE-1 | GCTCCAGGTCCACGAGGATGGCTCGA | D'afallaa, T. |
| 3451 DBM-B2-5RACE-2 | CAATCTGGTCCCGCACTGG | D'afallaa, T. |
| 3602 Hr5BbsF | acgcgaagacggcgccgtttaaaattgaactggcttacg | D'afallaa, T. |
| 3603 Hr5BbsR | TGCGCGAACACTGCTGCAAATGAATTATTTAATTATCAATC | D'afallaa, T. |
| 3604 TubProBbs F | acgcgaagacataaggcactgcagtgaacggttatgatgg | D'afallaa, T. |
| 3605 Tub5utrBbs R | TCGCGAACACTATGGTCATTTGGTAGTTGAGTTTACAAATTACAC | D'afallaa, T. |
| 3606 TcTraBsmb F | acgccgtctcacaaaaatgttaagttcttagattaataaattatc | D'afallaa, T. |
| 3607 TcTraBsmb R | ACGCCGTCTCACTCATCTGAAATCAAGCGGTTCTC | D'afallaa, T. |
| 3608 DsRBsmbF | agcgcgttcagaaggctagaatggcctccgaa | D'afallaa, T. |
| 3609 DsRBsmbR | TCGCGTCTCGCTAGCTTACAGGAACAGGTGGTGGCGG | D'afallaa, T. |
| 3610 Diag1TcTra | AAGCATTATGGACGTGACGTTAG | D'afallaa, T. |
| 3630 CG8121E4R | GAAGAACTCAGACTGGTGGCCTT | D'afallaa, T. |
| 3631 CG8121E5R | GTTCATCTGCTCCGAAACG | D'afallaa, T. |
| 3632 CG9222E3R | GTCGATGTGCTGCTCCTTGC | D'afallaa, T. |
| 3633 DBMtNr6E F | CTACATCAGAAACATACCCAGAGAAC | D'afallaa, T. |
| 3634DBMtNr6ER | CGGGCACAGTAGAGTTACGTG | D'afallaa, T. |
| 3635DBMtNr6E7R | GTGTCGCAATTGTGGACATC | D'afallaa, T. |
| 3636DBMtNr6E9R | CGGTTTCGATTCCAGGTAGG | D'afallaa, T. |
| 3637DBMB2tubE1 F | CCTGTCTAGTAGGTTGAAGTGAGCTG | D'afallaa, T. |
| 3638DBMB2tubE2 R | CTCGAATTCGAGATCAGTAGCG | D'afallaa, T. |
| 3639DBMtN17E1F | ATGTCTGGTTGCTGATTTAGTTA | D'afallaa, T. |
| 3640DBMtN17E3 R | CAGCAGTAGTACAAGTAGCACTTCTTCC | D'afallaa, T. |
| 3658B2TubKpnF | tcagggaggtaaccggaaaattcggaa | D'afallaa, T. |
| 3659B2TubBsmbR | TGCGCGTCTCGATATCACTTCCCAGAACTGAAATG | D'afallaa, T. |
| 3660 DsRBsmbF | agcgcgtcttatcgccctccgagaacgtcatc | D'afallaa, T. |
| 3661 SV40RsrR | ACGCTCGGACCGGGTCCAGACATGATAAGATAACATTGATG | D'afallaa, T. |
| 3662 Diag1B2Tu b | ggaggtttcggccatc | D'afallaa, T. |

| | | | |
|------|----------------|---|---------------|
| 3658 | DBM-B2 | TCAGGGAGGTACCGAAAATTCGGAA | D'afallaa, T. |
| 3692 | GFP | ccttgtagatgaaggcagccgtcctgca | D'afallaa, T. |
| 3702 | DBM-B2TubF | cctgtcttagttaggtgaagtggatcgagg | D'afallaa, T. |
| 3703 | B2TubBsmBR | AGCGCGTCTCACGATATCACTTCCCAGAACTGAAATGTAAC | D'afallaa, T. |
| 3704 | UbiBsmB-F | agcgcgtctctatcgcttaggcagatcttcgtcaagaccctgac | D'afallaa, T. |
| 3705 | UbiBsmB-R | AGCGCGTCTCTACCACCGCCAGGCCAG | D'afallaa, T. |
| 3706 | DsR-BsmBF | agcgcgtctcatggggcgcccccccccaagaagaag | D'afallaa, T. |
| 3707 | DsR-BsmBR | AGCGCGTCTCAGATCTGGGCCCTACGGGTCCACCTCCG | D'afallaa, T. |
| 3708 | DmProtAsc F | acatggcgcccaagttcaaataatgttaatggatgcagg | D'afallaa, T. |
| 3709 | DsR-ApaR | CTATGGGCCCCAGGAACACAGGTGGTGGCGGC | D'afallaa, T. |
| 3832 | TcTub1BsmbF | agcgcgtctcaggcccactgcgtgaacg | D'afallaa, T. |
| 3833 | DsRBsmbR | AGCGCGTCTCACACGGGCCGTCGGAGGGGA | D'afallaa, T. |
| 3834 | TcTraBsmb F | agcgcgtctcacgtggtaattgggtttcatagtcttgacga | D'afallaa, T. |
| 3835 | TcTraBsmb R | AGCGCGTCTCAGCATCTGAAATCAAGCGTTCTCTCA | D'afallaa, T. |
| 3836 | DsRBsmbF | agcgcgtcttatgcagaagaagaccatggctg | D'afallaa, T. |
| 3837 | SV40BsmbR | AGCGCGTCTCACGCCGTTAAGATACTTGATGAGTTGGA C | D'afallaa, T. |
| 3739 | UbiBbsF | acgcgaagacacccatggcagatcttcgtcaagacc | D'afallaa, T. |
| 3845 | TcFRNAi-F1 | cgtctaatacgactcaactatagggagaacgttattaccactcaccgagac | D'afallaa, T. |
| 3846 | TcFRNAi-R1 | cgtctaatacgactcaactatagggagaGATACATCATCGGATGCCTC AA | D'afallaa, T. |
| 3847 | TcFRNAi-F2 | cgtctaatacgactcaactatagggagaagacgttagcagatcgccgag | D'afallaa, T. |
| 3848 | TcFRNAi-R2 | cgtctaatacgactcaactatagggagaCCGGTTGGTACGCGATTTC A | D'afallaa, T. |
| 3849 | TcMRNAi-F | cgtctaatacgactcaactatagggagaataattggtttcatagtcttgacgaag | D'afallaa, T. |
| 3850 | TcMRNAi-R | cgtctaatacgactcaactatagggagaTTTGATGAAATTCTGA CTGTTTTATGC | D'afallaa, T. |
| 3976 | dbmB2tubE1f | GCCAGTGCAGAACAGATTGGA | D'afallaa, T. |
| 3977 | dbmB2tubaxR | cttccttcctcatcaaactcgc | D'afallaa, T. |
| 3978 | dbmB2tubp1R | caaaaccacttacaaaaacatagaatacc | D'afallaa, T. |
| 3979 | dbmB2tubp2R | tcatgttttattaaaacaaaaccacttac | D'afallaa, T. |
| 3980 | dbmB2tubp3R | cctatgtattcactctactataatactg | D'afallaa, T. |
| 3981 | dbmB2tubp4R | gtgcaaaatctaaataataccatcagc | D'afallaa, T. |
| 3994 | aTuProBsmBF | agcgcgtctcgccactgcgtgaacggttatgtgg | D'afallaa, T. |
| 3995 | ZsGBsmBF | AGCGCGTCTCCATGCCCTGCCGGTGATCAC | D'afallaa, T. |
| 3996 | TcTraBsmb F | agcgcgtctcaggatggtaattgggtttcatagtcttgacg | D'afallaa, T. |

| | | |
|--------------------------------|--|---------------|
| 3997 TctraBsmB R | ACGCGCGTCTCATCTGCATCTGAAATCAAGCGGTTCTCTC | D'afallaa, T. |
| 3998 DsRBsmBF | agcgctcgtccatggcttcttcgtcatcac | D'afallaa, T. |
| 3999 DsRBsmBR | AGCGCGTCTCACTGCCACCATGGGTACCGCTAGAGTCG | D'afallaa, T. |
| 4000hsp70BsmBR | agcgctcgtccatggcttcttcgtcatcac | D'afallaa, T. |
| 4001 teto7BsmB F | ACGCGCGTCTCAGAGATCTGCTAGCCAGGTTTCGACT | D'afallaa, T. |
| 4002 hsp70BbsF | acgcgaagacagtctcagcgccgaggataaatagaggc | D'afallaa, T. |
| 4003 SV40BbsR | TGCGCGAAGACGCCGGCCGATCATAATCAGCCATACCACATTGTAG | D'afallaa, T. |
| 4004 UbiqBbsF | acgcgaagacatcagatctcgtaagaccctgacc | D'afallaa, T. |
| 4005 midDsRBbs R | CGTGATGCAGAAGAACCATGGGCTG | D'afallaa, T. |
| 4006 TcTraIntF1 | tgcgtacttctgttagttctccgat | D'afallaa, T. |
| 4007 TcTraIntF2 | catcaaatgagagaacccgttg | D'afallaa, T. |
| 4040 tTAVBbsR | TCGCGAAGACATCCTACCCACCGTACTCGTCAATTCCA | D'afallaa, T. |
| 4041 B2TubBbsF | acgcgaagacactaggcgttagtatttgtattgttacatgg | D'afallaa, T. |
| 4042 B2TubBbsR | TCGCGAAGACATGAGTGTGCAAAATCTAAATAATACCATCAGC | D'afallaa, T. |
| 4043 IE1BbsF | acgcgaagacagctagcatgccgtaacggacc | D'afallaa, T. |
| 4044 SV40BbsR | TCGCGAAGACTGACTCGCGTTAAGATACTTGATGAGTTG | D'afallaa, T. |
| 4045 DiagB2Tub | gttcgctgcgctaattcacaag | D'afallaa, T. |
| PBW beta2tub motif rev | CCTTCCTCCTTCTCGTCAAACCTCTCCTCCTCCTC | Harris, C. |
| PBW beta2 for across intron | ATCGGAGCAAAGTTCTGGGAAGTAATATC | Harris, C. |
| DBM beta2 for across intron | ATTGGAGCTAAGTTCTGGGAAGTGATATC | Harris, C. |
| Prot intron seqF | CGCATCTCAATGAATTTCACAGAACTG | Harris, C. |
| Prot intron SeqR | CCATCCGACATCGAGACAAACAC | Harris, C. |

9.4 TRANSFORMATION EFFICIENCY USING PIGGYBAC LITERATURE REVIEW

The literature review in Chapter 2 produced a dataframe of studies and the associated microinjection and transformation data (Table 54).

Table 54. A literature search of “piggyBac AND insect*” was manually searched for publications with instances of successful transformation of an insect species using *piggyBac*. Publications were read and examined for relevant details. NA means not available, information was not provided in the paper or supplementary material.

| Order | Species | Embryo s injected | Fertile injection survivor s G ₀ | Success X | Success rate X/ G ₀ | Reference |
|-------------|----------------------------|-------------------|---|-----------|--------------------------------|--------------------------|
| Lepidoptera | <i>Bombyx mori</i> | 1058 | 230 | 3 | 0.013 | Tamura et al., 2000 |
| Lepidoptera | <i>Bombyx mori</i> | 1440 | 424 | 9 | 0.021 | Tamura et al., 2000 |
| Lepidoptera | <i>Plutella xylostella</i> | 3104 | 584 | 4 | 0.0068 | Martins et al., 2012 |
| Lepidoptera | <i>Plutella xylostella</i> | 1925 | 1462 | 7 | 0.0047 | Martins et al., 2012 |
| Lepidoptera | <i>Plutella xylostella</i> | 2410 | 922 | 6 | 0.0065 | Martins et al., 2012 |
| Lepidoptera | <i>Plutella xylostella</i> | 1202 | 463 | 3 | 0.0064 | Martins et al., 2012 |
| Diptera | <i>Anopheles stephensi</i> | 337 | NA | 5 | NA | Catteruccia et al., 2005 |
| Diptera | <i>Anastrepha ludens</i> | NA | 247 | 12 | 0.048 | Condon et al., 2007 |
| Diptera | <i>Anastrepha ludens</i> | NA | 132 | 12 | 0.091 | Condon et al., 2008 |
| Diptera | <i>Ceratitis capitata</i> | NA | NA | NA | NA | Gong et al., 2005 |
| Diptera | <i>Aedes aegypti</i> | 2000 | 225 | 3 | 0.013 | Kokoza et al., 2001 |
| Diptera | <i>Aedes albopictus</i> | 6000 | 550 | 6 | 0.011 | Labbe et al., 2010 |
| Diptera | <i>Aedes albopictus</i> | 604 | 32 | 1 | 0.031 | Labbe et al., 2010 |
| Diptera | <i>Aedes albopictus</i> | 2052 | 86 | 1 | 0.012 | Labbe et al., 2010 |
| Diptera | <i>Aedes albopictus</i> | 2165 | 161 | 2 | 0.012 | Labbe et al., 2010 |
| Coleoptera | <i>Tribolium castaneum</i> | 509 | 146 | 61 | 0.42 | Lorenzen et al., 2003 |
| Coleoptera | <i>Tribolium castaneum</i> | 314 | 95 | 38 | 0.40 | Lorenzen et al., 2003 |
| Coleoptera | <i>Tribolium castaneum</i> | 459 | 152 | 36 | 0.24 | Lorenzen et al., 2003 |
| Coleoptera | <i>Tribolium castaneum</i> | 612 | 50 | 12 | 0.24 | Lorenzen et al., 2003 |

| | | | | | | |
|-------------|---------------------------------|------|-----|----|-------|---------------------------|
| Coleoptera | <i>Tribolium castaneum</i> | 392 | 55 | 12 | 0.22 | Lorenzen et al., 2003 |
| Lepidoptera | <i>Pectinophora gossypiella</i> | 5974 | 86 | 3 | 0.035 | Peloquin et al., 2000 |
| Lepidoptera | <i>Pectinophora gossypiella</i> | 1911 | 0 | 0 | 0 | Peloquin et al., 2000 |
| Diptera | <i>Ceratitis capitata</i> | 500 | 135 | 2 | 0.015 | Morrison et al., 2009 |
| Diptera | <i>Drosophila suzukii</i> | 75 | 33 | 4 | 0.12 | Schetelig & Handler, 2013 |
| Lepidoptera | <i>Bombyx mori</i> | 5000 | 367 | 18 | 0.049 | Zhuang et al., 2010 |
| Lepidoptera | <i>Bombyx mori</i> | 2600 | 32 | 0 | 0 | Zhuang et al., 2010 |
| Lepidoptera | <i>Bombyx mori</i> | 2400 | 27 | 5 | 0.19 | Zhuang et al., 2010 |
| Hymenoptera | <i>Athalia rosae</i> | 278 | 57 | 3 | 0.053 | Sumitani et al., 2003 |
| Diptera | <i>Anastrepha suspensa</i> | NA | NA | NA | NA | Schetelig & Handler, 2012 |
| Lepidoptera | <i>Ostrinia furnacalis</i> | NA | NA | NA | NA | Liu et al., 2012 |
| Diptera | <i>Bactrocera tryoni</i> | 155 | 19 | 2 | 0.11 | Raphael et al., 2011 |
| Diptera | <i>Bactrocera tryoni</i> | 465 | 39 | 4 | 0.10 | Raphael et al., 2011 |
| Diptera | <i>Bactrocera tryoni</i> | 595 | 36 | 9 | 0.25 | Raphael et al., 2011 |
| Diptera | <i>Bactrocera tryoni</i> | 723 | 13 | 0 | 0 | Raphael et al., 2011 |
| Diptera | <i>Bactrocera tryoni</i> | 640 | 25 | 0 | 0 | Raphael et al., 2011 |
| Diptera | <i>Lucilia cuprina</i> | 3000 | 66 | 4 | 0.061 | Concha et al., 2011 |
| Diptera | <i>Lucilia cuprina</i> | 3000 | 52 | 11 | 0.21 | Concha et al., 2011 |
| Diptera | <i>Teleopsis dalmanni</i> | 699 | 50 | 2 | 0.04 | Warren et al., 2010 |
| Diptera | <i>Drosophila melanogaster</i> | 2650 | 283 | 4 | 0.014 | Handler & Harrell, 1999 |
| Diptera | <i>Drosophila melanogaster</i> | 1940 | 122 | 11 | 0.090 | Handler & Harrell, 1999 |
| Diptera | <i>Drosophila melanogaster</i> | 2147 | 218 | 7 | 0.032 | Handler & Harrell, 1999 |
| Diptera | <i>Musca domestica</i> | 1668 | 41 | 7 | 0.17 | Hediger et al., 2001 |

| | | | | | | |
|------------|--------------------------------|------|-----|----|--------|-----------------------|
| Coleoptera | <i>Tribolium castaneum</i> | 700 | 69 | 3 | 0.043 | Lorenzen et al., 2002 |
| Diptera | <i>Anopheles albimanus</i> | 1715 | 35 | 12 | 0.35 | Perera et al., 2002 |
| Diptera | <i>Anopheles gambiae</i> | 1954 | 172 | 1 | 0.0058 | Grossman et al., 2001 |
| Diptera | <i>Lucilia cuprina</i> | 800 | 27 | 0 | 0 | Heinrich et al., 2002 |
| Diptera | <i>Lucilia cuprina</i> | 800 | 29 | 0 | 0 | Heinrich et al., 2002 |
| Diptera | <i>Lucilia cuprina</i> | 2400 | 121 | 2 | 0.017 | Heinrich et al., 2002 |
| Diptera | <i>Cochliomyia hominivorax</i> | 2180 | 22 | 3 | 0.14 | Allen et al., 2004 |
| Diptera | <i>Drosophila simulans</i> | 1680 | 16 | 2 | 0.13 | Depra et al., 2004 |
| Coleoptera | <i>Harmonia axyridis</i> | 405 | 54 | 2 | 0.037 | Kuwayama et al., 2006 |
| Diptera | <i>Drosophila melanogaster</i> | 2730 | 166 | 1 | 0.0060 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 930 | 70 | 6 | 0.086 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 620 | 16 | 2 | 0.13 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 650 | 20 | 3 | 0.15 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 730 | 31 | 4 | 0.13 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 670 | 28 | 3 | 0.11 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 710 | 31 | 3 | 0.097 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 850 | 36 | 5 | 0.14 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 990 | 86 | 0 | 0 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 610 | 71 | 0 | 0 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 840 | 69 | 0 | 0 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 650 | 31 | 4 | 0.13 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 530 | 36 | 5 | 0.14 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 1020 | 36 | 8 | 0.22 | Li et al, 2005 |

| | | | | | | |
|---------|--------------------------------|-----|----|----|------|-----------------------|
| Diptera | <i>Drosophila melanogaster</i> | 515 | 22 | 8 | 0.36 | Li et al, 2005 |
| Diptera | <i>Drosophila melanogaster</i> | 533 | 88 | 22 | 0.25 | Li et al, 2005 |
| Diptera | <i>Drosophila willistoni</i> | 539 | 6 | 4 | 0.66 | Finokiet et al., 2007 |

This table is available on request as a .csv file. Data was manipulated and analysed in R.

9.5 PROTEIN ASSAY FOR DSRED IN OX5133

The transcript was shown to be found at similar levels relative to the endogenous gene which suggests the transcript number is not the limitation. Protein from OX5116 and OX5133 were examined for DsRed protein with a positive control. A positive of DsRed expressing in the germline would have provided the ideal control, however we lacked this and considered using either a germline DsRed-expressing insect from another species or a diamondback moth line showing somatic expression of DsRed.

9.5.1 A WESTERN BLOTH WITH A MONOCLONAL PRIMARY ANTIBODY FOR DSRED.

We opted for the former, to use the OX4700 *Tribolium* line which shows DsRed expression as its transgenesis marker as the positive control (ten female pupae). A bacteria-expressed DsRed was also included as a positive control for the antibody (recombinant *Disco* DsRed2 Protein, Clontech Laboratories, US). Testes were removed from 10 males of OX5133 and homogenised with a mortar and pestle. A Western blot for DsRed was carried out. An additional sample included two pooled OX5145 female pupae (in which we had found DsRed transcription but no fluorescence).

The testes of the diamondback moth transgenic lines were likely to contain the associated transgenesis marker protein, as elucidated by fluorescence microscopy. This was AmCyan for OX5133 and DsRed for OX4026. Dissection of the testes inevitably brought with it somatic tissue and the associated fluorescent protein (Figure 167). Wild-type testes were negative for fluorescence (not shown).

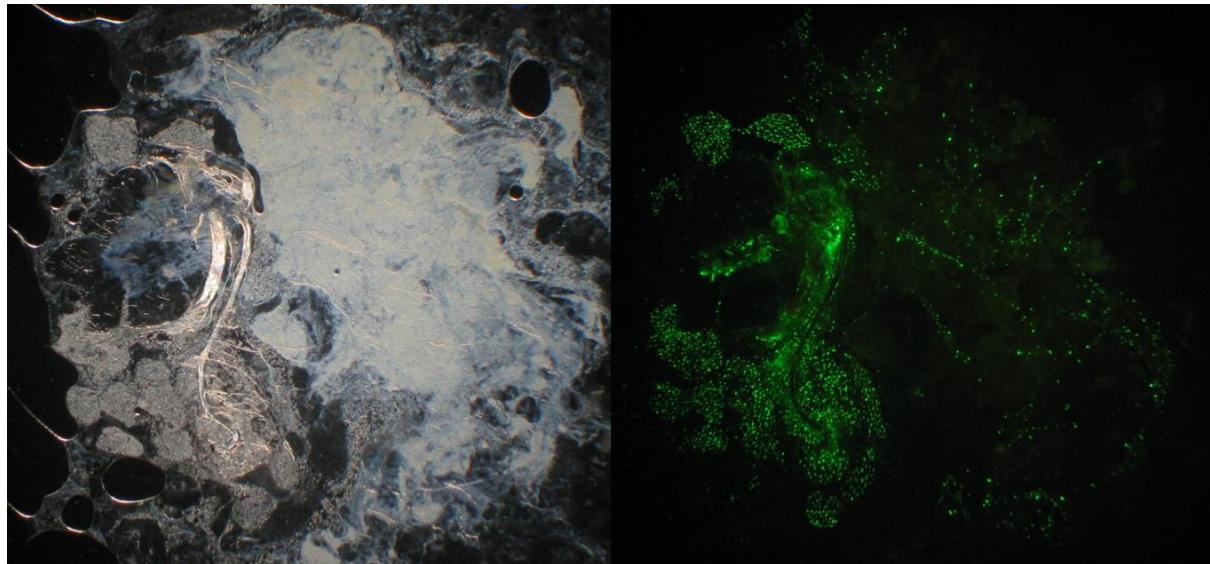


Figure 167. 2-day-old adult testes smear of OX5133 under white and blue light. The correct filter was not available for AmCyan so green filter was used which resulted in detection by bleed through. The nuclei can be seen due to the AmCyan being localised to the nucleus. A typical testis has a diameter of 700 µm, the dimensions and features are distorted due to the weight of the glass slide. DsRed was not detectable. The testes were dissected and pooled for protein immunoblotting.

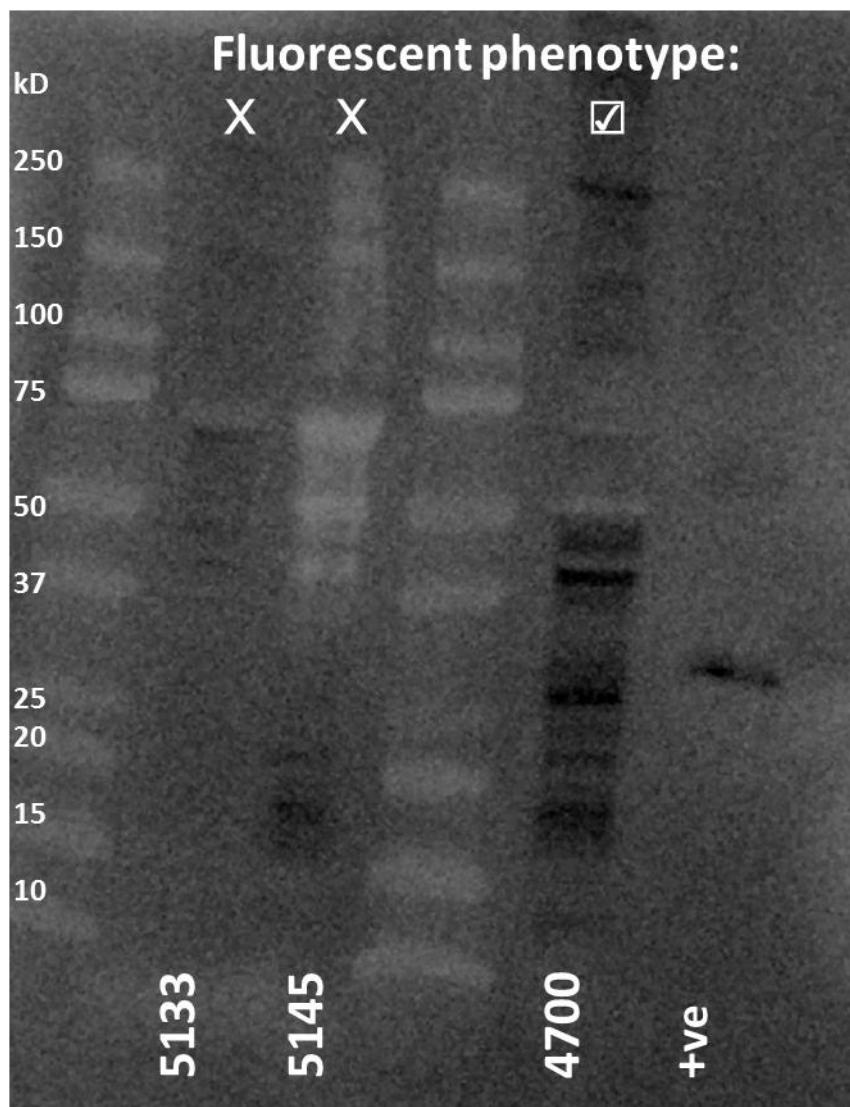


Figure 168. A Western blot with a monoclonal primary antibody for DsRed. Protein imaged on a membrane using Chemi detection. It shows how those insect lines with DsRed transcript but no fluorescence have no detectable DsRed protein. The DsRed-positive OX4700 show some specific staining albeit at different sizes compared to the positive control of pure DsRed2. Molecular masses in kDa are indicated to the left of a protein ladder.

The outcome of the Western blot was unclear, with banding of a different size in our positive *in vivo* control. No antibody staining of DsRed was evident in the non-fluorescent OX5145 and OX5133 insect lines. Anti-body staining indicated by black band formation was detected in the DsRed fluorescent OX4700 line albeit not at the same position as the positive control.

The positive control DsRed2 was made by *Escherichia coli*, and comprises an additional His tag, which may account for the larger size relative to our DsRed-positive insect line OX4700. The quality control data of the product (Catalogue number: 632436) states that “*A major band of 30-35 kDa was observed. Though DsRed2 has a calculated mass of ~25.7 kDa, it usually runs above 29 kDa on SDS-polyacrylamide gels. A minor band, representing a possible cleavage product, with an apparent molecular mass of about 22 kDa, may also be observed under some conditions.*”

The other bands could be dimer and tetramer versions of DsRed but do not appear to be exactly 2× and 4× the size.

Comparison to DsRed cleavage in the literature shows that the heat treatment at 95 °C for 5 min may have been ineffective in denaturing the larger ~ 200 kDa band shown in C (Gross *et al.*, 2000). The cleavage product at 22 kDa may be present in our OX4700 assay. The expected banding is not entirely inconsistent with our *in vivo* OX4700 DsRed staining.

The monoclonal antibody was repeated with diamondback moth samples including a positive control for diamondback moth (OX4026) and *Tribolium* DsRed-expressing OX4700C. A negative control of wild-type diamondback moth was included for comparison. Reduced sample volume was also loaded to avoid saturation with 20 µl of testes sample, 10 µl of upper half of the diamondback moth carcass and half the recommended positive control to reduce saturation of signal. Only 5 µl of the wild-type carcass upper was used as preliminary work had shown the protein concentration to be higher than in the other samples. Quantification was unnecessary as we were interested in a qualitative answer to the presence or absence of DsRed.

Tribolium OX4700C may be reliable as a positive control as it is around the expected ~ 26 kDa corroborating Figure 168. *Tribolium* OX4700C was the only DsRed fluorescent line that tested positive for putative DsRed protein. All the diamondback moth samples were negative for a ~ 26 kDa band but instead showed banding between 15 -20 kDa suggesting cross reactivity.

We did not detect DsRed in our diamondback moth positive control OX4026 therefore we cannot conclude that no DsRed was translated in OX5133 carcass or testes. Concerning the negative controls, the wild-type testes were positive at the band between 15-20 kDa shared with the other diamondback moth except for the upper carcass. This may have been due to insufficient sample being used as it was half that of the other diamondback moth "u" samples.

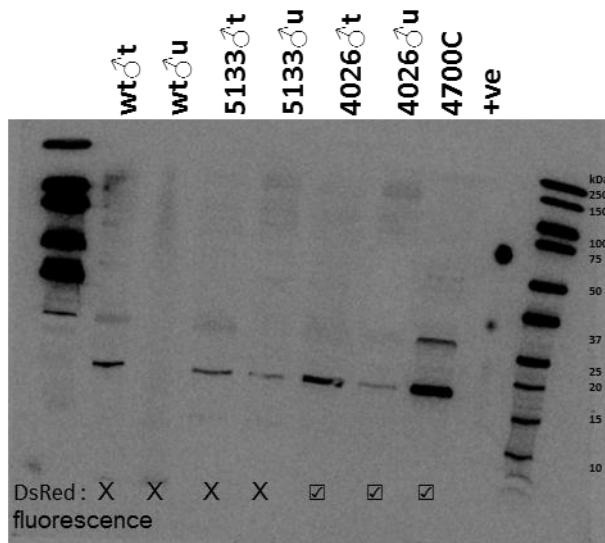


Figure 169. A Western blot with a monoclonal primary antibody for DsRed. The positive control of bacterially produced DsRed is circular due to squashing of the gel during transfer. The position of the bands can still be seen as circles at the expected sizes. An additional positive control of bacterially synthesised DsRed was also included, the his-tag increases the size of this band relative to the expected size of *in vivo* DsRed produced by insects. Wt (wild-type), t suffix for testes, u for upper non-testes carcass. OX5133 has cyan as the transgenesis marker and DsRed promoted by diamondback moth β 2-tubulin, OX4026 had DsRed as the main transgenesis marker.

9.5.2 A WESTERN BLOTH WITH A POLYCLONAL PRIMARY ANTIBODY FOR DSRED.

OX4700 females were used for comparison with OX5145 females, controlling for sex, but marker reliability assays (Chapter 5) suggest that the females produce much less DsRed than the males, perhaps a male is a better beetle DsRed-positive control.

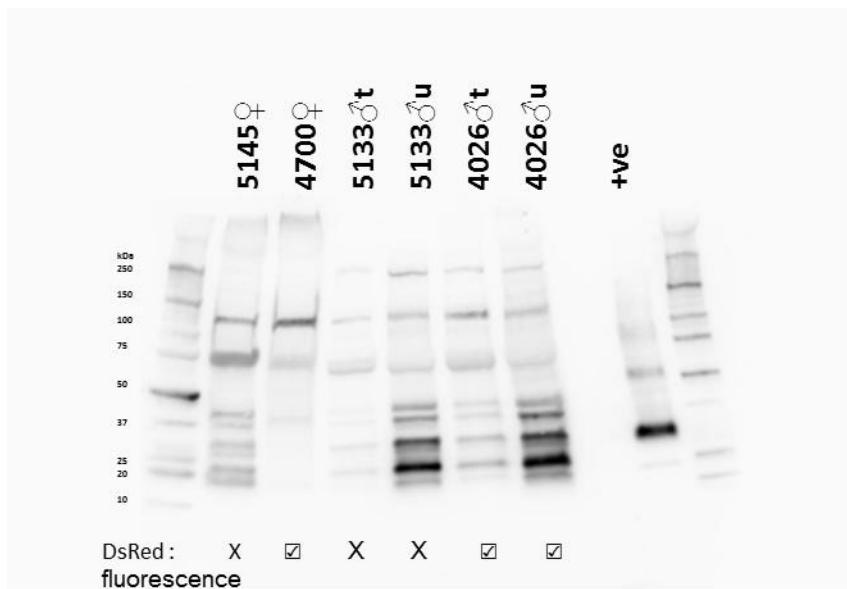
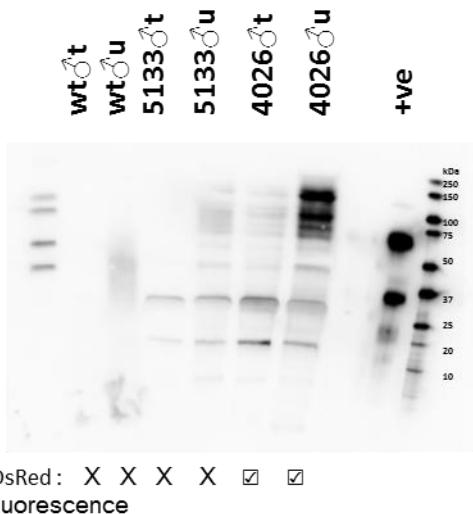


Figure 170. A Western blot with a polyclonal primary antibody for DsRed. There is a lot of cross reactivity with the antibody hitting multiple epitopes probably non-specific to DsRed. An additional positive control of bacterially synthesised DsRed was also included, the his-tag increases the size of this band relative to the expected size of *in vivo* DsRed produced by insects. Wt (wild-type), t suffix for testes, u for upper non-testes carcass. OX5133 has cyan as the transgenesis marker and DsRed promoted by diamondback moth $\beta 2$ -tubulin, OX4026 had DsRed as the main transgenesis marker.

This is lacking a negative control and it also confuses with the inclusion of two species. The blot was repeated with wild-type negative control for diamondback moth. In this blot the negatives remain blank with no banding.

Unlike the monoclonal the polyclonal suggests the presence of DsRed in OX5133, sharing similar banding to the DsRed fluorescent line OX4026. However, if DsRed shares epitopes with AmCyan perhaps we are seeing cross-reactivity, the antibody binding to a non-intended target.



DsRed: X X X X
fluorescence

Figure 171. A Western blot with a polyclonal primary antibody for DsRed with negative controls of wild-type diamondback moth and a positive control of a DsRed containing diamondback moth line OX4026. An additional positive control of bacterially synthesised DsRed was also included, the histag increases the size of this band relative to the expected size of *in vivo* DsRed produced by insects. Wt (wild-type), t suffix for testes, u for upper non-testes carcass. OX5133 has cyan as the transgenesis marker and DsRed promoted by diamondback moth β 2 tubulin, OX4026 had DsRed as the main transgenesis marker.

BLASTing the protein sequence of AmCyan against DsRed revealed some similarity between the proteins possibly providing common epitopes for the polyclonal antibody to hit. This seems likely as the polyclonal suggested DsRed was in the upper abdomen of OX5133 but we know from our RT-PCR that there is no (or very little) transcript in this part of the insect male.

| Score 196 bits(498) | Expect 3e-66 | Method Compositional matrix adjust. | Identities 95/204(47%) | Positives 131/204(64%) | Gaps 2/204(0%) |
|------------------------|---|--|---------------------------|---------------------------|-------------------|
| Query 1 | MASSENVITEFMRFKVVRMEGTGNGHEFEIEGEGEGRPYEGHNTVKLKVT- | -KGGPLPFAN | 58 | | |
| Sbjct 17 | MA S I M+ M+G VIGH F ++GEG G+PYEG T KVT GGPL F++ | | | | |
| | MALSNKFIDDDMKHTYHMDGCVIGHYFTVKGEESGKPYEGTQTSTFKVTHANGPLAFSF | | 76 | | |
| Query 59 | DILSPQFQYGSKVVKHPADIPDKKLSPFEGFKWERNVNNFEDGGVATVTQDSSLQDGCF | 118 | | | |
| Sbjct 77 | DILS F YG++ +P +PDY K +FP+G +ER +EDGGVAT + SL+ CF | | | | |
| | DILSTVFMYGNRCFTAYPTSMPPDFKQAFPDGHNSYERTFTYEDGGVATASHEISLKGNCF | 136 | | | |
| Query 119 | IYKVKFIGVNFPSDGPWVQKKTMGWEASTERLYPRDGVLKGETHKALKLKDGHHYLVEFK | 178 | | | |
| Sbjct 137 | +K F GVNFPA+DGPWM KKT GN+ S E++ DG+LKG+ L L+ GG+Y +F | | | | |
| | EHKSTFHGVNFPA+DGPWM+AKTTGNDPSFEKMTVCDGILKGDVTAFLMLQGGGNYRCQFH | 196 | | | |
| Query 179 | SIYMAKKPVQLPGYYYYDAKLDIT 202 | | | | |
| Sbjct 197 | + Y KKPV +P + V+ ++ T | | | | |
| | TSYKTKKPVTPNPHVVEHRIART 220 | | | | |

9.6 OPTIMISING *TRIBOLIUM CASTANEUM* MICRO-INJECTIONS FOR INSECT TRANSGENESIS

9.6.1 ISSUES FOR MASS-REARING AND ADEQUATE EMBRYO PRODUCTION FOR INJECTION

9.6.1.1 GENERAL ISSUES

To develop RIDL technology in the species it will be necessary for all life stages to ingest tetracycline; however they live in and feed on a dry medium, a variation from other insects in which similar work has been conducted.

9.6.1.2 DEVELOPMENT

Tribolium is straightforward to rear and can complete its life-cycle under a variety of temperature-humidity combinations (Figure 172). However, as insect transgenesis is limited by generation time as well as egg production, one should endeavor to provide optimal conditions for both egg-to-adult development time, as well as female fecundity (i.e. net eggs produced). Although the strains and laboratory conditions are different, comparison of several different samples of this type by a range of authors corroborated the conclusions inferred from the figure below (Sokoloff, 1974). Howe's investigations also showed that the sex ratio was 1:1 for all conditions. Visible light has no effect on the development of *Tribolium* (Sokoloff, 1974) but can deter oviposition (Hawk *et al*, 1972).

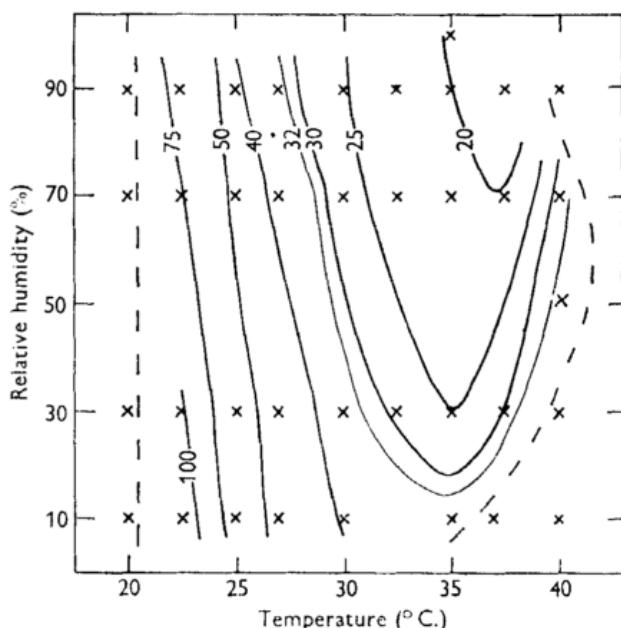


Figure 172. The periods in days required for complete development from egg-laying to the formation of the adult of *Tribolium castaneum* at various combinations of temperature and relative humidity. Crosses mark the temperature-humidity combinations for which the results were obtained from which the positions of the lines were determined. Full lines join all conditions at which development is completed in the number of days stated. Dotted lines show limits beyond which development cannot be completed. Taken from Howe (1956).

When rearing *Tribolium* on standard medium (whole wheat flour and brewer's yeast, 95:5 by mass) the species averages six moults at the larval stage (Mickel and Standish, 1947). The exuvia can interfere with sieving and the egg collection and de-chorionation process. It is important to develop a method to separate healthy beetles from exuvia, the dead or dying beetles to avoid disease in the colony.

9.6.1.3 OVIPOSITION

Oviposition in *Tribolium* is characterised by long-lived imagos which produce eggs continuously over a long period. This attribute, combined with a reasonably brief life-cycle, has helped to make *Tribolium* a widely studied model organism. Using the associated background knowledge to establish how to maximise oviposition and net egg production will be important in generating enough eggs for microinjection and an insect transgenesis programme. Howe (1962) provides a variety of concerns and caveats when interpreting oviposition data as any experimental manipulation may affect the measurable outcome.

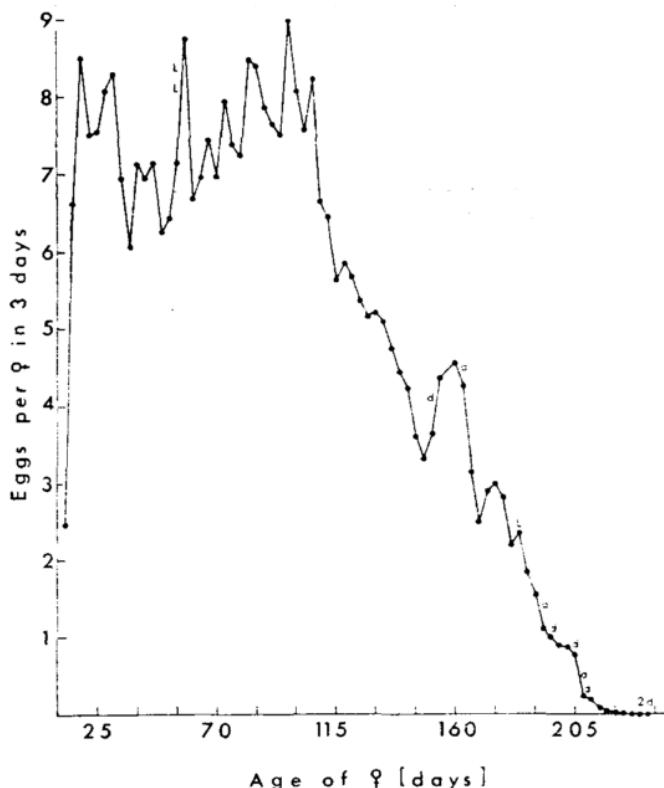


Figure 173. The average three-daily oviposition rate of females of *Tribolium castaneum* at 25°C and 70%relative humidity over the whole adult life. L indicates that a female was lost, and d that one died, at the times shown. Taken from Howe 1962a.

Females start to oviposit at a minimum of 4 days and a maximum up to an indefinite period after eclosion (with the maximum dependent on environmental conditions). Male ability to inseminate remains constant throughout life, but female fecundity (specifically egg laying rate) begins to decline after 3-4 months. Oviposition is positively correlated with temperature and relative humidity – with a temperature relative-humidity interaction evident (Howe, 1962a).

9.6.1.4 TYPE OF CONTAINER

Howe observed that oviposition is depressed by insect density and the shape of the container used. It is unclear as to what style of container is optimum for oviposition, although it is suggested that females prefer to burrow prior to oviposition.

9.6.1.5 CANNIBALISM

As a model organism, one major disadvantageous biological attribute of *Tribolium* is that of inter-stage predation. Upon oviposition into the medium, an egg is at risk of predation – in the scenario of producing eggs for microinjection, the egg is at risk from adults only.

Rich (1956) carried out a series of experiments to determine how insect density affects oviposition and cannibalism. His experimental design consisted of marking 200 eggs and mixing them in with flour and brewer's yeast medium (95:5 by mass (8 g total) and adding beetles. He recorded the change in count of marked and unmarked eggs over time.

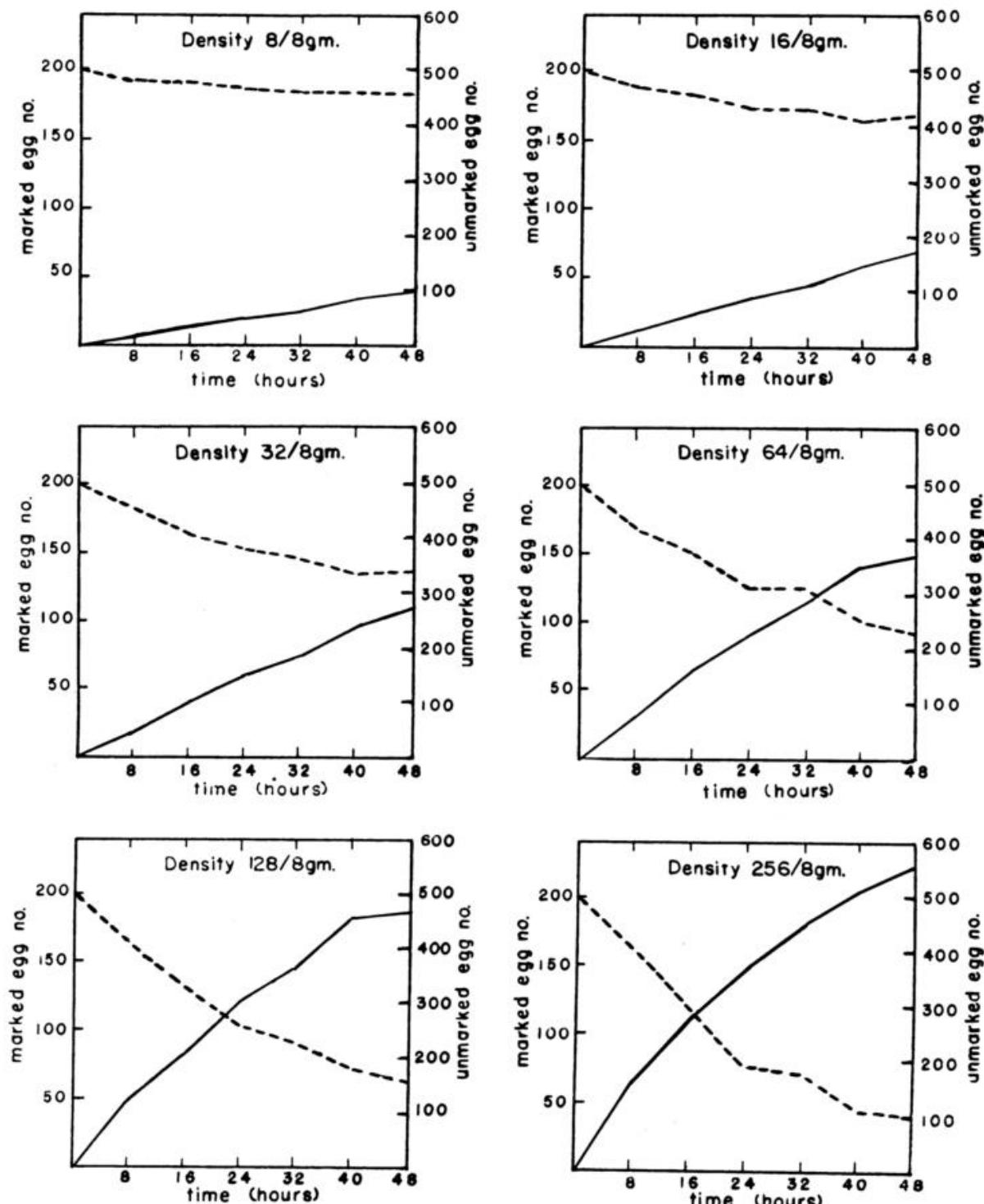


Figure 174. Number of eggs recovered relative to each time interval for each density (broken line - marked eggs; solid line – unmarked eggs). Taken from Rich (1956).

The gradient of the marked egg number is steeper with increasing density. At each density the decline in marked egg number is fairly linear over time. The increase in unmarked eggs follows an inverse pattern. The summation of both lines gives the overall number of eggs present, at all densities this increases over time.

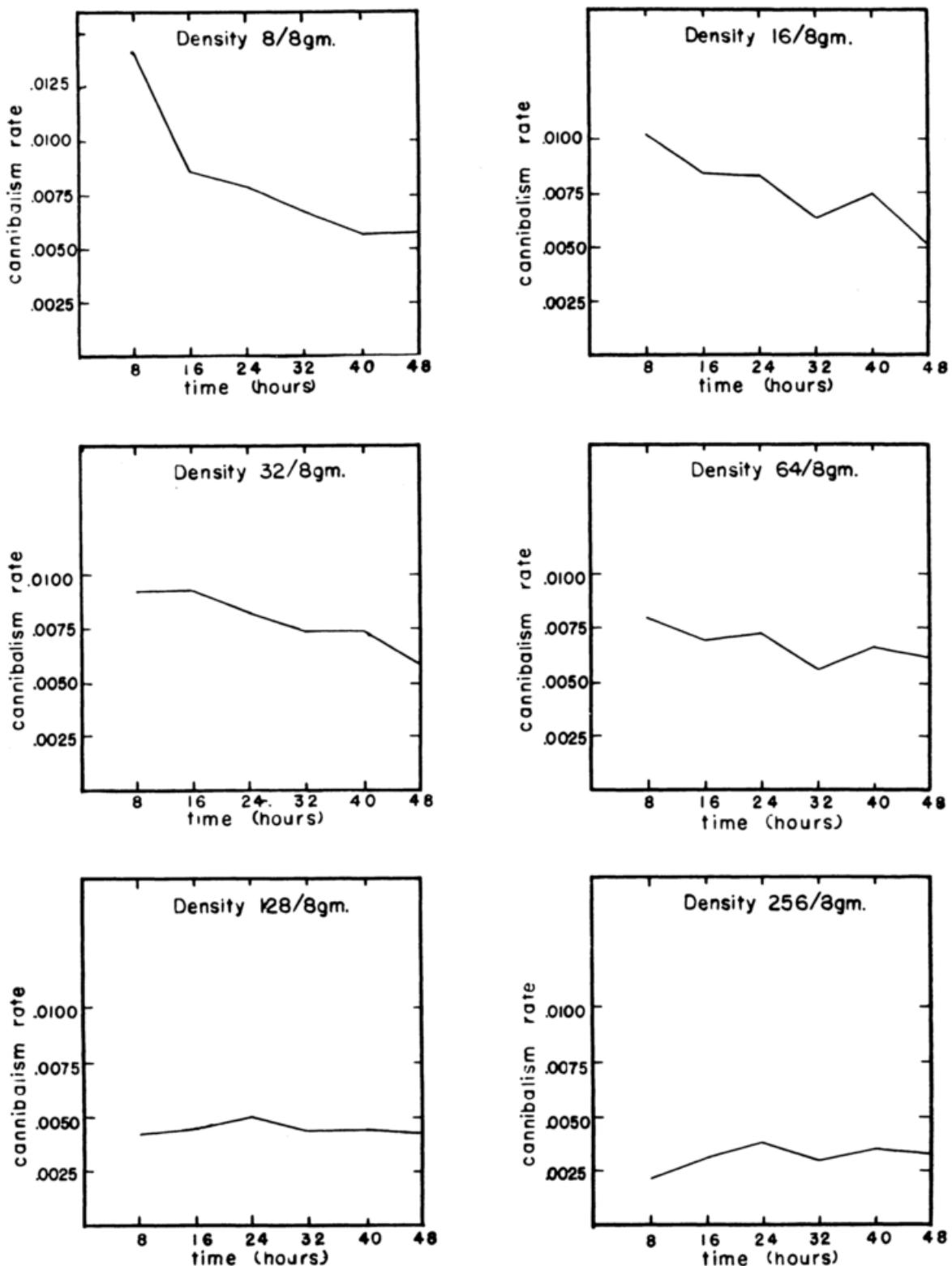


Figure 175. Cannibalism rate relative to time interval. Cannibalism rate = the fraction of eggs present which are cannibalised per beetle in unit time. Taken from Rich (1956).

Both the cannibalism rate (Figure 175) and fecundity rate (Figure 176) are not constant with density, between time intervals or within time interval, between densities. As density increases the table and graphs show these relations to hold:

- i. cannibalism as a rate per beetle decreases
- ii. real fecundity, again as a rate per beetle, decreases
- iii. the change in gradient is similar for cannibalism and real fecundity rates as the density increases (Rich, 1956).

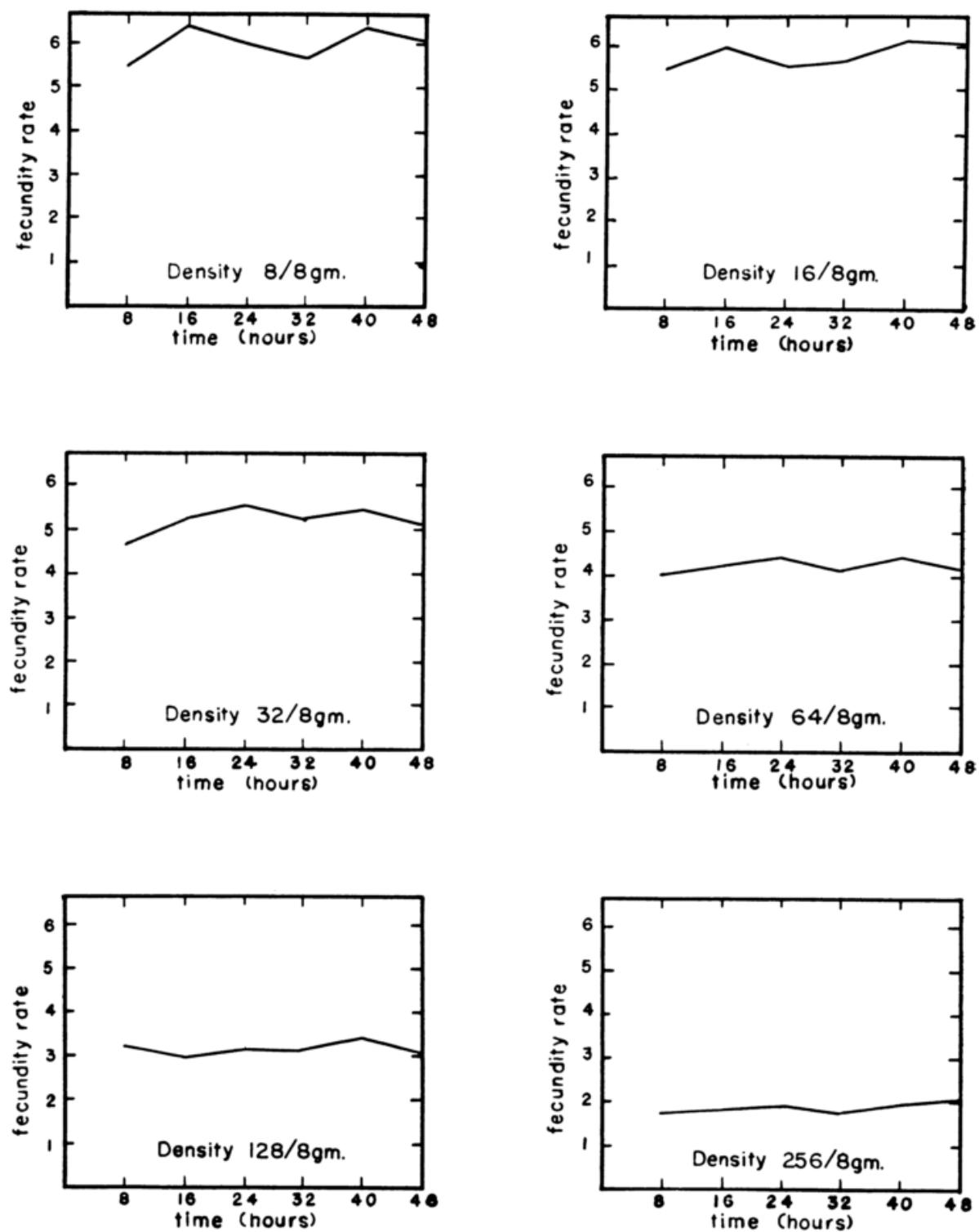


Figure 176. Real Fecundity rate relative to time interval. Real fecundity rate = the number of eggs produced per beetle in unit time. Taken from Rich (1956).

Although these experiments do not look at the first two hours of egg-laying they provide a starting point for interpolation. At these densities we could calculate the expected number of eggs produced in two hours by using the mean eight hour real fecundity rate.

Table 55. Number of adults per 8 g of medium and the associated fecundity rate. The expected number of eggs is the product of number of adults and the mean fecundity rate, calculated from Rich (1956).

| Density (adults / 8 g) | Mean Fecundity Rate | Expected number of eggs per unit (0 d.p) |
|-------------------------------|----------------------------|---|
| 8 | 5.4 | 43 |
| 16 | 5.5 | 88 |
| 32 | 4.65 | 148 |
| 64 | 4.03 | 258 |
| 128 | 3.2 | 410 |
| 256 | 2.1 | 538 |

From these calculations it seems that, despite the trade-off between instances of cannibalism and frequency of copulation, a higher density will provide a greater number of eggs for injection. Of course this is based heavily upon potentially unrealistic assumptions and does not consider other factors, such as time taken for beetles to acclimatise to the new environment and recover from the disturbance associated with handling.

It should be emphasised that egg cannibalism is consistently present and is a feature of *Tribolium*, however that does not mean it cannot be mitigated against, with intensity of egg cannibalism in the literature shown to vary from 15% to 76% (Sokoloff, 1972).

Oviposition is also affected by the medium – the quantity and quality of flour. In a no-choice bioassay experiment, individual females were placed in an arena with varying amounts of flour (Campbell & Runnion, 2003). Approximately 0.2 g of flour was the threshold for maximum number of eggs laid (Figure 177). However as flour is not a limiting factor in our methodology (flour is cheap) an excess of this would be recommended as density affects are likely to complicate matters. In choice chamber experiments *Tribolium* has been shown to prefer a flour layer of > 7 mm depth compared to a 0.5 mm-deep flour layer (Stanley and Grundmann (1965) quoted in Sokoloff (1972)).

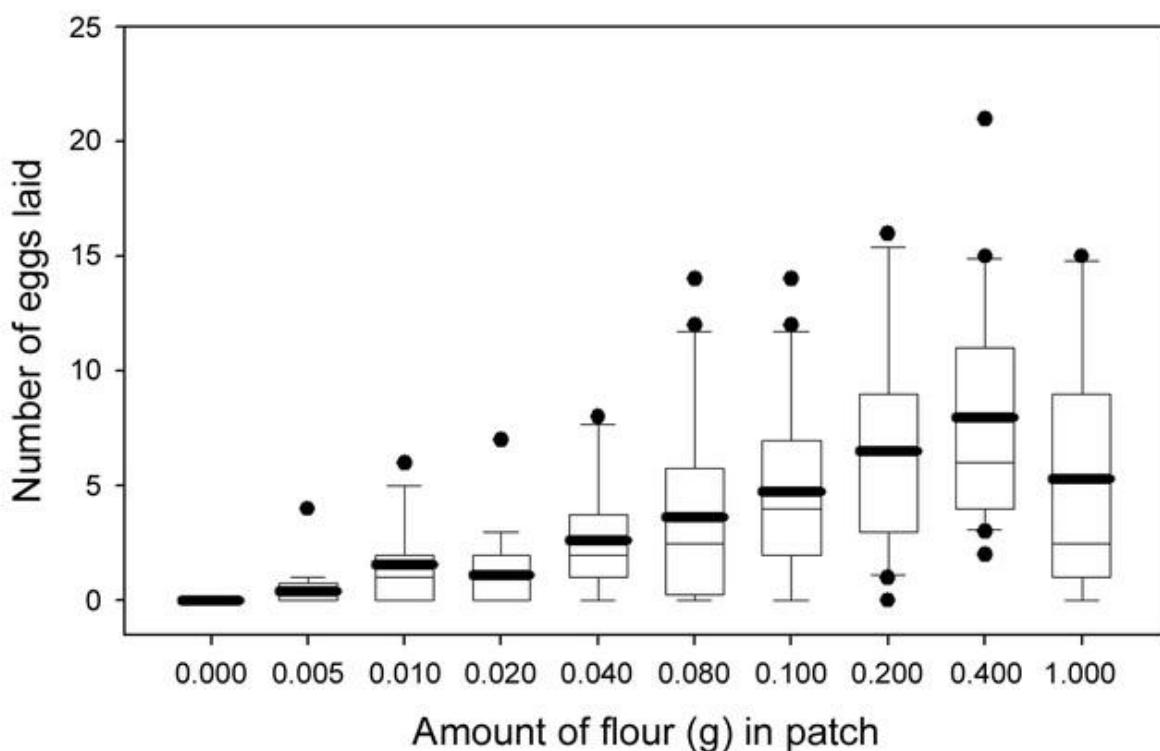


Figure 177. The number of eggs laid by female *Tribolium castaneum* in different amounts of flour in a no-choice experiment. Boxes represent 25th to 75th percentile of the data, whiskers represent the 10th and 90th percentiles, circles represent outliers, thin lines through box represent the median and thick lines represent the mean. Taken from Campbell and Runnion (2003).

Female survival can be affected by a high-density mating environment at standard temperatures (Grazer & Martin, 2011); this can be mitigated against by increasing the temperature (from 30°C to 34°C).

The conclusions from the literature reviewed above, combined with the recommendations of Berghammer *et al.*, (2009), allowed the design of a standard operating procedure for germline transformation in the red flour beetle found in the methods section, with inevitable refinement to local conditions and restrictions throughout the programme of study.

9.6.2 EMBRYO MASS PRODUCTION TROUBLESHOOTING

9.6.2.1 PROCESS OPTIMISATION

Following methodology in the literature there were some issues with embryo production due to conflicting advice (Park, 1934; Pavlopoulos *et al.*, 2004; Berghammer *et al.*, 2009). Egg production from 10 g of beetles was consistently under the expected number of approximately 210 eggs (3*70), given a 2-h egg lay, with average egg collection below 20 per egg lay. This was unsatisfactory for insect transgenesis as a large number of embryos are required, considering that egg hatch is reduced after injection and less than 1% of injection survivors may yield transgenic progeny. A variety of factors were investigated using controlled experiments to determine effect on embryo count.

In step 10 of the method described by Berghammer (2009), the author suggests to, “Separate eggs from flour using a 300- μm sieve”. However, using this sieve size it was observed that, of the embryos that were found, many were stuck in the holes pole-end first. Due to the eggs’ flexible nature and the vigorous sieving it seemed probable that many of the eggs were simply passing through the mesh (embryo size has been shown to vary due to differences between particle size of the growing medium (Park, 1934). An alternative 250- μm sieve was used. Results improved dramatically with a return much closer to the expected number (350 eggs, expected 270).

Another modification to the method involved avoiding the use of egg baskets. The author discovered that following the sieving step the *Tribolium* embryos could be observed on the sieve surface when placed under a microscope. The eggs could simply be transferred from sieve to glass slide by a fine brush, lined up ready for microinjection. Dechorionation was also found to be unnecessary given appropriate needle design and moistening the embryo pole prior to penetration.

9.6.3 EMBRYO COLLECTION

The control mortality for egg collection in *Tribolium* was determined from Berghammer methodology (2010). This was pertinent for *Tribolium* as the embryos are exposed to greater stress (the chemical stress of de-chorionation and the mechanical stress of sieving and brushing) when extracted from the flour ready for injection compared to the lepidopteran embryos, thus providing information to adapt the model in Chapter 5. Slides with eggs were stored in two separate Tupperware boxes which acted as replicates. A significant difference revealed by a test of equal proportions in survival occurred between the boxes, despite identical treatments ($X_2 = 8.12$, $df = 1$, $p = 0.0044$).

Mean survival was in the range $0.46 < 0.51 < 0.56$ (two-sided 95% confidence interval determined from a beta-posterior, using “binom.bayes” function in R, $n = 358$.) (The choice of method or approximation used to generate the 95% confidence interval is not as important with the mean nearer to 0.5, towards the bounds of zero or one the Clopper-Pearson method has shown to be more robust in simulations (Winkler *et al.*, 2002; Ludbrook & Mew, 2009)).

This procedure is to be repeated in order to produce a less variable measure of control survival. The control survival is expected to be lower than the true survival of an embryo laid in flour, as the embryo must be extracted from the flour, manipulated mechanically and counted. The control treatment should be indicative of the survival of an embryo exposed to identical treatment but not injected for transgenesis.

9.6.4 EMBRYO CARE

9.6.4.1 EXPERIMENTAL DESIGN

A preliminary experiment was designed to elucidate factors that influence the survival of injected embryos. Temperature was assumed to be constant (incubator measured $32 \pm 2^\circ\text{C}$ with two types of thermometer). The experiment was set up as shown with a blocked design. Embryos were collected as described in the Methods chapter. Treatments were distributed evenly, to correct for any order effects the embryos (as sets of five) were assigned to slides in a

randomised order. The replicate was at the level of the Petri dish (if one ignores that the experiment took place in only one incubator). Each replicate consisted of a vector of the success or failure of five pseudo-replicates (the development of an egg into a larva) (Figure 178).

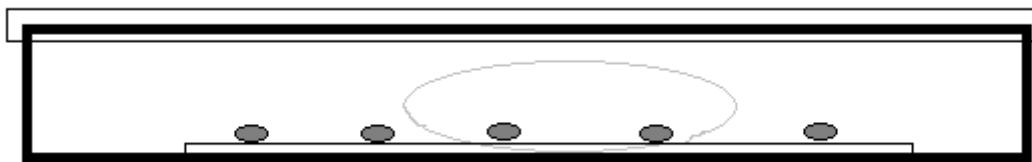


Figure 178. The experimental unit – a Petri dish was either sealed with Parafilm or it was not. All Petri dishes had cotton wool of equivalent size placed on one side saturated with distilled water. A microscope slide with five embryos that had either been injected or not injected was placed into the centre of the dish (punctured with a capillary). The number of eggs that hatched or did not hatch were counted and combined into a vectorised response variable after a 4-day incubation period. The presence or absence of condensate was noted. Not drawn to scale.

It is important to understand the factors that are associated with an embryo developing to larval stage or not. The experimental design permits us to know how many times this event occurred and how many times it did not occur, given certain conditions; a response vector was constructed.

The hypothesis that there is no interaction between the use of Parafilm (Parafilm used or not used) and injection status (stabbed with a capillary or not stabbed) on the survival rate of the embryos to larvae. A generalised linear model was used with a factorial analysis of the factors given above. Analysis of deviance with binomial errors was appropriate for the categorical explanatory variables.

9.6.4.2 RESULTS

The model was examined for over-dispersion; the residual deviance is 19.016 on 16 degrees of freedom, an over-dispersion factor of 1.189. The interaction term was non-significant ($p > 0.95$). It was removed from the model and an F-test was used to compare the original and simplified models. The next step was to see if any further model simplification was possible, an analysis of deviance revealed a significant difference between the injection status on embryo survival, but it is not obvious that Parafilm status was needed in the model so it was removed. The lower the p value the lower our confidence in the null hypothesis being true.

Table 56. The minimal adequate model output or summary of a glm with binomial error distribution of the response variable (alive, dead). This model was further simplified by removal of the Parafilm status term, an F test revealed no significant loss in explanatory power.

| Factor | Residual Df | Residual Deviance | F | Pr (>F) |
|-----------------|-------------|-------------------|------|---------|
| Parafilm status | 18 | 34.13 | 0.91 | 0.34 |
| Injection | 17 | 24.32 | 9.81 | 0.002 |

| | | | | |
|--------|--|--|--|--|
| status | | | | |
|--------|--|--|--|--|

The minimal adequate model contains just two parameters; the mean survival rates of the embryos when not injected and injected. The generalised linear model produced the correct mean proportions (the means were back-transformed from logits) for each level of injection status (Control and Injected).

Table 57. The two parameters required to adequately describe the data, the mean survival rate of the embryos to first instar larvae given the two levels of injected or un-injected (control).

| Group | Mean survival rate to larvae | Standard error of the mean |
|----------|------------------------------|----------------------------|
| Control | 0.320 | 0.0660 |
| Injected | 0.082 | 0.0384 |

A two-sample test for equality of proportions with continuity correction was also performed. The two mean proportions were significantly different from one another ($\chi^2 = 7.6$ df = 1, p = 0.006).

9.6.4.3 CONCLUSION

This preliminary experiment highlights the high mortality associated with embryo collection and preparation for microinjection, compounded further by the associated “stabbing” reducing the hatch rate to 10%.

The data suggest low survival may impede the efforts to generate a transgenic line in *Tribolium*.

9.6.5 POST-INJECTION SURVIVAL OF EMBRYOS

The literature suggests the storage of embryos in a humidified chamber post-injection improves hatch rate. Embryos were injected while on glass slides, each slide acted as a replicate with over thirty eggs per slide. The derived variable hatch rate was calculated from the number of larvae successfully emerging from injected embryos (number of trials). Glass slide were placed into an airtight tupperware box two sheets of absorbent paper (30 cm x 10 cm) placed onto the bottom under the slides. The tissue paper was saturated with 20 ml of treatment solution. After two days chambers were opened and glass slides were placed into Petri dishes. The Tupperware chamber was resealed with the Petri dishes inside. Hatch rate was checked after three, four and five days to ensure all embryos had sufficient time to hatch.

No injected embryos survived to larvae when treated with a dry Tupperware chamber (30 °C and 40% relative humidity).

Three treatments of interest were tap water (hereafter water, 30 °C and 95% relative humidity), 15% (by mass) NaCl solution (30 °C and 90% relative humidity) and 30% NaCl solution 30 °C and (80% relative humidity). Measurement of the relative humidity with a datalogger revealed real humidity differences between the treatments.

The non-Gaussian variation between replicates made the differences in mean survival difficult to interpret from inspection of summary statistics. Graphical inspection of the distribution of mean survival was preferred using a cumulative rank scatterplot (empirical cumulative distribution function) due to the differing number of replicates per treatment. Although not often used this type of distribution display is superior to a histogram as it does not suffer from arbitrary bin width assignment. The relative rank represents the proportion of observations less than or equal to the x axis variable, in this case the hatch rate. Unlike Tukey's box and whisker all data points are provided for each distribution. A cumulative rank scatterplot sampled from the standard normal distribution is provided for reference in Figure 180.

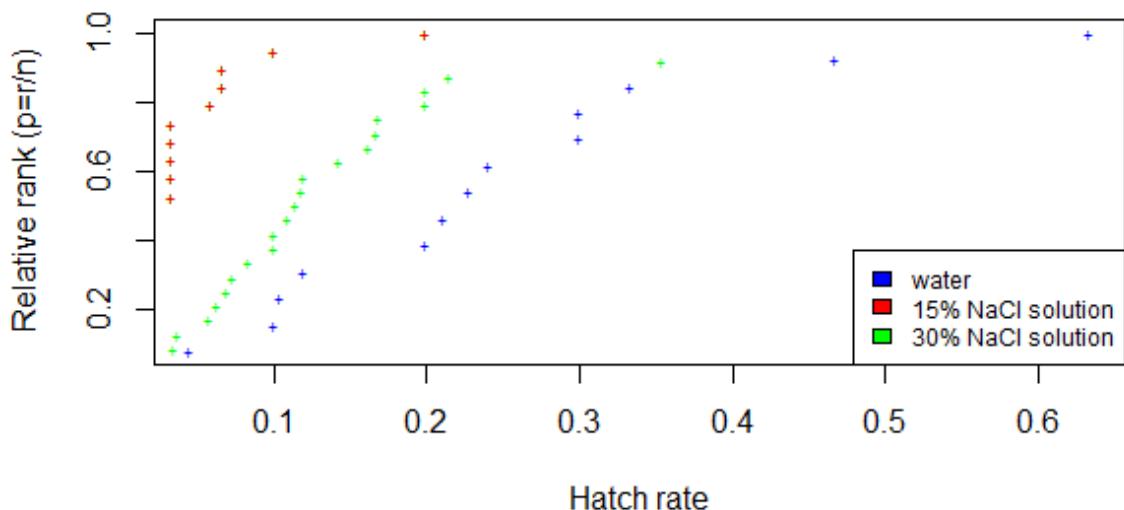


Figure 179. Cumulative rank scatterplot of the empirical probability mass functions of the different treatments correlated with different conditions in the Tupperware humidified chamber (water - 95% relative humidity; 15% NaCl – 90% relative humidity; 30% NaCl – 80% relative humidity). All chambers were stored in an incubator at 30 °C. The relative rank describes the proportion of observations equal to or less than the value read off the x axis. Code: tc_injects.R

Comparison of medians (relative rank = 0.5) for each treatment reveals those curves shifted to the right had a superior hatch rate. All treatments had a minimum at near zero. The shallower the gradient of the treatment lines the larger the variation between points within that treatment.

The distribution of water hatch rates is most sigmoidal and hence normal although it is still far from being appropriate for a Gaussian model or parametric testing between treatments. Due to variation in embryos per replicate additional noise may arise from the data. Accordingly statistical analysis was not undertaken. The embryo collection and injection process was monitored following this experiment and it was suggested that 30% NaCl soaked tissues provide the best characterised humidified chamber for increasing hatch rate. However, further replicates for water should be run alongside future injections for improved resolution of the analysis.

The results were surprising in that the hatch rate associated with the treatments did not increase with humidity; the 30% NaCl solution was superior to the 15 % NaCl solution with the 0% NaCl solution superior to both. The increase in humidity provided a trade-off; although tap water provided greatest humidity and consequent survival it also introduced mould problems which represented a health risk as well as difficulties in removing adult beetles due to mycelium growth. A compromise of 5% NaCl solution was used for the first four days after injection until hatching. Larvae containing pots where stored in a humid chamber with paper towels soaked in 30% NaCl solution to prevent mould.

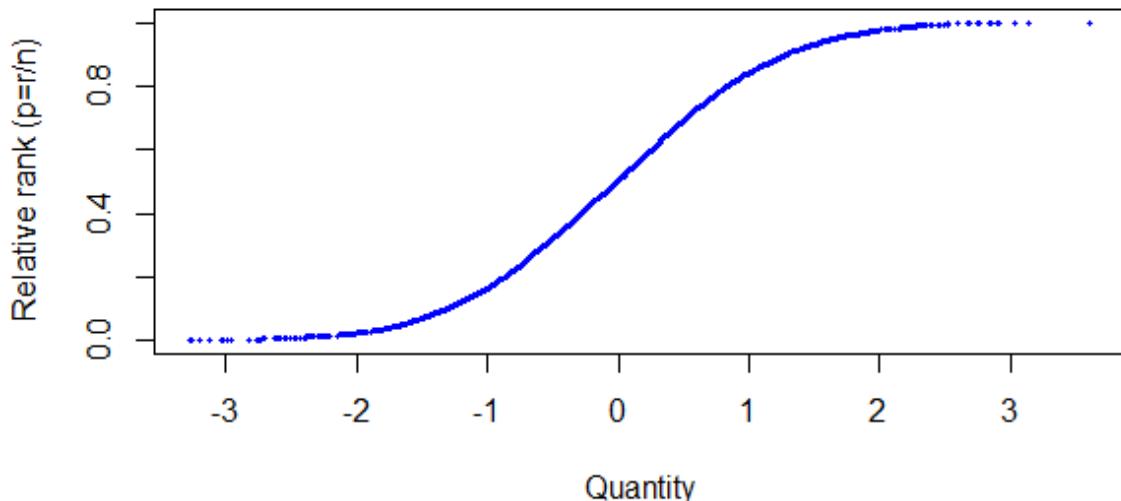


Figure 180. A cumulative rank scatterplot of 3000 random samples of the standard normal distribution (`set.seed(255)` in R, using the function `rnorm(3000)`). A demonstration of the appearance of a normally distributed variable using the cumulative rank scatterplot. Code: `tc_injects.R`

9.6.6 POST-SIEVE EMBRYO TREATMENT PRIOR TO INJECTION

9.6.6.1 THREE DAYS HUMID ONE DAY DRY EMBRYO TREATMENT

Embryos laid by beetles will be collected following the default method (See Chapter 6). The post bleach treatment (factor with two levels: wet; embryos were moved from the sieve using a brush into a pool of tap water on the glass slide, and dry; embryos will be moved directly from the sieve using a brush onto the glass slide) effect on embryo survival to hatch was investigated.

Ten embryos were placed on each slide with the treatment alternately assigned to remove order effects. When the last embryo was placed onto the slide it was then placed into a plastic Tupperware box with tap-water-saturated tissue paper lining the bottom to increase humidity. After three days the slides were placed into a dry Tupperware box to remove the risk of fungal

mycelium related mortality and drowning of wandering larvae in water droplets. A day later the hatch rate was investigated.

Each slide acted as an experimental unit with the ten pseudo-replicate embryos per slide. The numbers of successes (survival or mortality of the embryo to first instar larvae) were recorded, with the number of failures described as 'n', 10 in all cases, minus the number of successes.

A 5-sample test for equality of proportions without continuity correction was first used to determine any significant differences between replicates for each treatment (dry; $X^2 = 8.0$, $df = 4$, $p = 0.09$, wet; $X^2 = 1.8$, $df = 4$, $p = 0.78$). As no statistical difference was found, $p > 0.05$, then pooling the data were justified (Xu *et al.*, 2010), with the acceptance of the null hypothesis that the five replicates from which the proportions were drawn have the same true proportion of embryo survival to first instar. The p value can be thought of as the strength of the evidence against the null hypothesis, the variability in the "wet" survival resulted in the smaller p value compared to the dry.

A chi-squared test was used on a two by two contingency table constructed from the observed number of successes (hatch) and failures (no hatch) between the two treatments. No significant difference between the two treatments was observed as represented visually by the 95% confidence intervals (Figure 181).

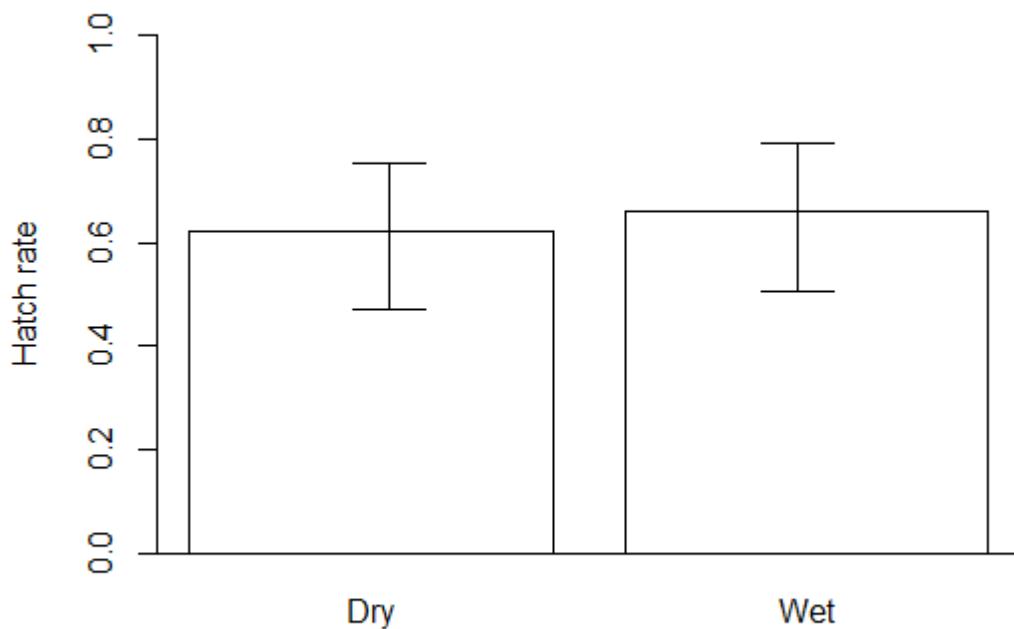


Figure 181. The hatch rate of *Tribolium castaneum* embryos exposed to two types of post dechorionation and egg sieving care. Embryos were removed from the sieve using a fine brush and placed onto a glass slide with (wet) or without (dry) a puddle of tap water. The mean proportion of embryos that gave rise to first instar larvae is given with 95% confidence intervals calculated using

the Clopper-Pearson method. No significant difference between the numbers of hatched and unhatched embryos between the treatments were observed ($X^2 = 0.0377$, df = 1, p = 0.85).

No difference between the two treatments was observed ($X^2 = 0.0377$, df = 1, p = 0.85). This is unsurprising as the embryos would still have some moisture from the submerged sieve, during the transfer of embryos to the slide only a small amount of evaporation would have taken place minimising the risk of dehydration to those embryos placed onto the slide using the dry method.

The dry method will be used in future microinjection as the remains of the flour in solution acts as an adhesive which can be used to fasten the eggs to the slide ready for injection.

A caveat, the statistical analyses above assume the replicate at the level of the embryo, when in reality the hatch rate of a given embryo is likely not independent of other embryos on the same slide.

9.6.6.2 THREE DAYS HUMID ONE DAY DRY EMBRYO TREATMENT WITH LED EXPOSURE

Perhaps a difference may be observed if the embryos are exposed to conditions more in line of those experienced by embryos during injection. Normally the slide would be on the microscope stand under an LED ring for several minutes (proportional to the number of embryos requiring injection).

Following the above experiment and the estimate of the mean survival of embryos post collection (p= 0.62), the number of replicates required to provide a given level of precision for a mean estimate was plotted to inform the experimental design and number of replicates carried out (Figure 182).

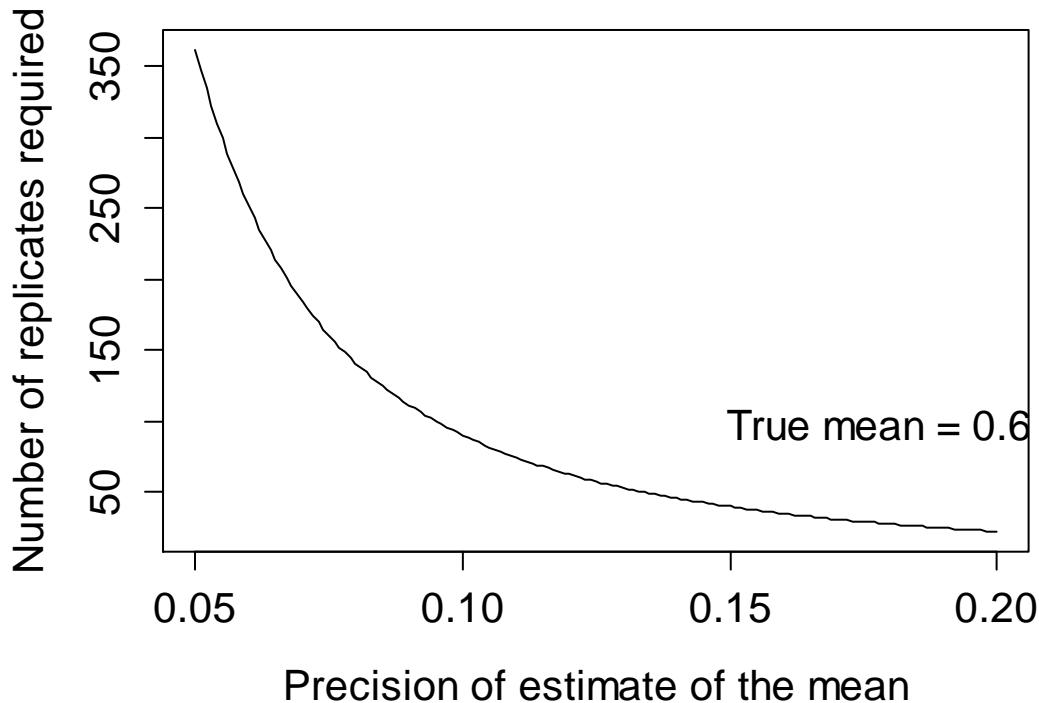


Figure 182. The number of replicates required to provide a given level of precision in statistic estimate with a mean survival of 0.62 using the formula, $1.96^2 * p * (1-p) / L^2$. For an estimate within 10% of the true mean, if the true mean were, $p = P = 0.62$ then 91 replicates would be required.

Based on the estimate of the proportion of embryo survival, the number of replicates required was rounded up to 100, to give a value within 10% of the true proportion.

The experiment was similar to that of 9.6.6.1 Three days humid one day dry embryo treatment, with difference in treatments applied. As no significant difference was found between the “wet” and “dry” treatment, the dry method was used as the control, as this benefits from the embryos being secured onto the slide for injection by the flour and water residue. The other treatment group was called “LED”, these embryos were exposed to a treatment where they were placed on the microscope stage, with a simulated injection procedure applied for twenty minutes (the slide was left on the stage and under an LED ring). The control slides were not exposed to this.

Ten embryos were placed on each slide with the treatment alternately assigned to remove order effects. When the last embryo was placed onto the slide it was exposed to LED treatment (or not for the control) then placed into a plastic Tupperware box with tap-water-saturated tissue paper lining the bottom to increase humidity. After three days the slides were placed into a separate Petri dish and again placed in a wet Tupperware box. Over the next two days hatch rate was assessed so that each embryo had been given sufficient time to hatch if it were going to.

Each slide acted as an experimental unit with the ten pseudo-replicate embryos per slide (as this was a pragmatic investigation where the methods reflected those that would be used each embryo could be considered a replicate. The numbers of successes (survival or mortality of the

embryo to first instar larvae) were recorded, with the number of failures described as 'n', 10 in all cases, minus the number of successes.

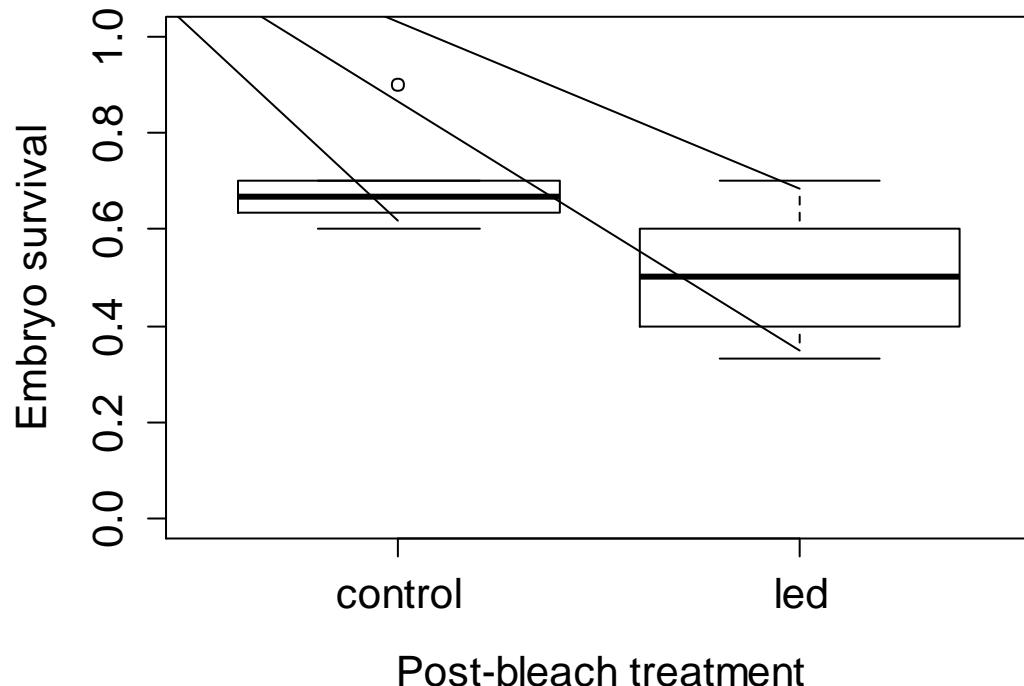


Figure 183. The distribution of the hatch rate proportion of embryos per slide of the two treatments. The bold line represents the median, the box the interquartile range and the whiskers the range of the data. The outlier is plotted if it exceed 1.5 times the interquartile range from the median.

A box and whisker shows the distribution of the per slide hatch rate of each treatment (Figure 183). The median survival of the control is larger than the embryos that were experienced putatively greater environmental stress. The "led" treatment survival data is more dispersed with a larger interquartile range and whiskers.

A 5-sample test for equality of proportions without continuity correction was first used to determine any significant differences between replicates for each treatment (control; $X^2 = 2.64$, $df = 4$, $p = 0.62$, LED; $X^2 = 2.99$, $df = 4$, $p = 0.56$). As no statistical difference was found, $p > 0.05$, then pooling the data were justified (Xu *et al.*, 2010), with the acceptance of the null hypothesis that the five replicates from which the proportions were drawn have the same true proportion of embryo survival to first instar. The p value can be thought of as the strength of the evidence against the null hypothesis.

A chi-squared test was used on a two by two contingency table constructed from the observed number of successes (hatch) and failures (no hatch) between the two treatments. No significant difference between the two treatments was observed ($X^2 = 2.51$, $df = 1$, $p = 0.11$) as represented visually by the 95% confidence intervals in Figure 183. The falling just short of statistical

significance can be thought of as less strong evidence for the null hypothesis, a greater number of embryos would have given superior power as described in Figure 182. The experiment should be repeated by greater sampling effort but is limited by the number of embryos produced per egg lay.

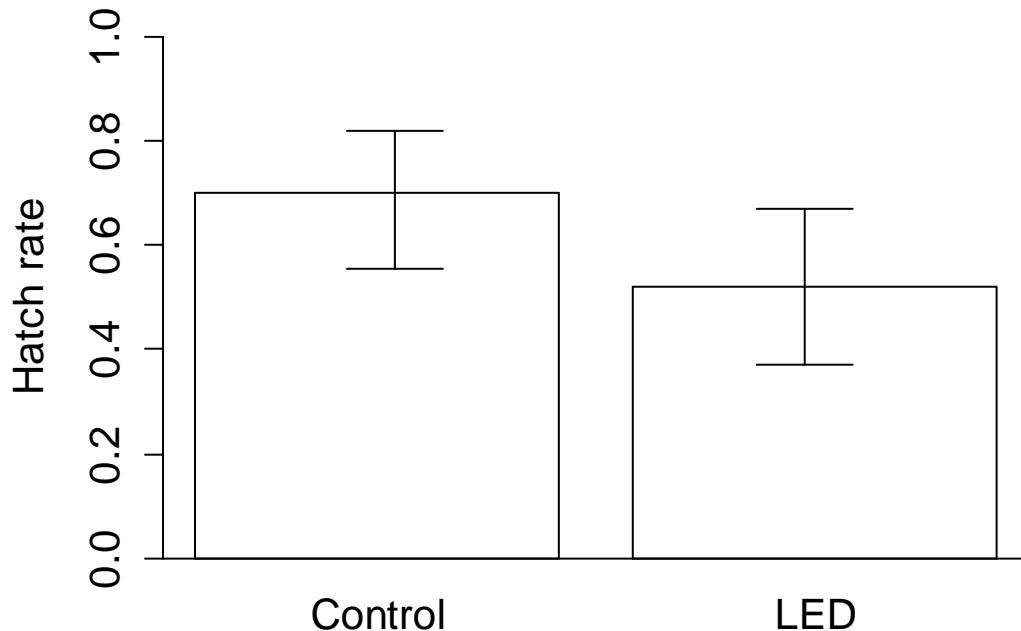


Figure 184. Mean hatch rate and 95% confidence intervals using the Clopper-Pearson method. The difference between the two means was not statistically significant ($X^2 = 2.51$, $df = 1$, $p = 0.11$).

There is weak evidence that the exposure of the dechorionated embryo to the environment does not reduce the hatch rate of the embryo ($p = 0.11$). It may be preferred to limit the number of embryos applied to each slide so that the embryos on the slide will spend less time exposed to a less humid and controlled environment. With other species it is typical to have hundreds of eggs on a slide for injection, however these have not been dechorionated, accordingly fewer embryos may be preferred. Monitoring of control mortality should follow.

The effect of injection compared to the control should be considered in detail next. This could be achieved using a similar methodology which would also allow comparison of mean hatch rates.

9.7 SUPPLEMENTARY METHODS FOR CHAPTER 2

9.7.1 LITERATURE REVIEW

9.7.1.1 DATA EXTRACTION AND QUALITY ASSESSMENT

Papers were checked for duplicates and removed as appropriate. Each included publication was read by MG and data extracted if it met the predefined criteria. If data were missing it was assigned a NA placeholder (NA – not available). The insect species, insect order, unique *piggyBac* construct ID (from the relevant paper to avoid duplication), number of embryos injected, injection survivors and independent transgenic lines derived from those injection survivor crosses were transliterated. The publication search and selection was repeated again 6 months after the initial study selection by the same reviewer and compared (March and October, 2013). Aside from two publications that were newly published, the second search found nine additional relevant publications, possibly due to a more systematic review approach and familiarity with the procedure.

The assumed publication bias will likely under represent the number of failures to transform, as publication of successful transformation of a novel species will tend to be preferred. Publications were inspected for pertinent information.

Most transformations with the *piggyBac* vector were achieved in this millennium. Eighty-seven different attempted or realised transformation events were described in the literature. Of these, the majority showed consensus in how to describe the outcomes of the experiment (75 provided complete information detailing injection number, G_0 and number of transgenic lines). Seven did not include any information about the numbers associated with transformation, asserting that it happened with molecular evidence as support.

9.7.1.1.2 TYPES OF OUTCOME MEASURES

Transformation efficiency was the preferred summary statistic of choice; the number of injection survivors (hereafter termed ' G_0 ') whose offspring were screened for successful transformation event (independent insertion event, hereafter termed 'X'). Transformation rate was not provided but could have been calculated from the data provided (number of injected embryos/X, discussed in detail in Adelman *et al.*, 2002; Martins *et al.*, 2013). Transformation efficiency was the preferred effect size measure (a standard measure by which all outcomes can be assessed) as it sidesteps the differential mortality rates between species after injection. If any embryo survives to adulthood following injection, what is the probability of it producing transgenic offspring?

9.7.1.1.3 BIAS CONSIDERATIONS

The biased nature of publications should be considered when interpreting data (Dubben & Beck-Bornholdt, 2005; Dwan *et al.*, 2013). Data describing a failure to achieve transformation, potentially considered less interesting by the research community, may be omitted from publications or rejected outright. This information, referred to as silent evidence, (Taleb, 2007) will produce a systemic error in our estimations of the likelihood that transgenesis will occur.

It is accepted that this review will be subject to publication bias. To what extent it is a problem is typically estimated using a funnel plot (Peters *et al.*, 2006). As the experiments become less precise (lower denominator or G_0 number), it would be expected for the results to be more variable. Interestingly a large number of experiments are close to zero and lie well outside the 99% confidence interval (assuming a binomial distribution, the CI is calculated using the Wilson method which has been shown to be less bad to other methods when dealing with extreme

probabilities (near zero or one)). These experiments showed great patience possibly injecting above and beyond what is typical. This may suggest publication bias and that the published experiments overestimate how difficult it is to achieve transformation using *piggyBac*.

The highest transformation efficiencies tend to be in those experiments with fewer number of injection survivors involved in crosses. However, using a mean of all insects injected with *piggyBac* assumes they are one population, the problem of this assumption is compounded by the proximity of the mean to zero which limits the variation at that bound.

In the least, the funnel plot (berryFunctions package in R by Berry Boessenkool (2014)) provides a caveat to those researchers too hasty to jump to conclusions about the cause of a deviation of an experiment's transformation efficiency from the mean; it is probably due to chance.

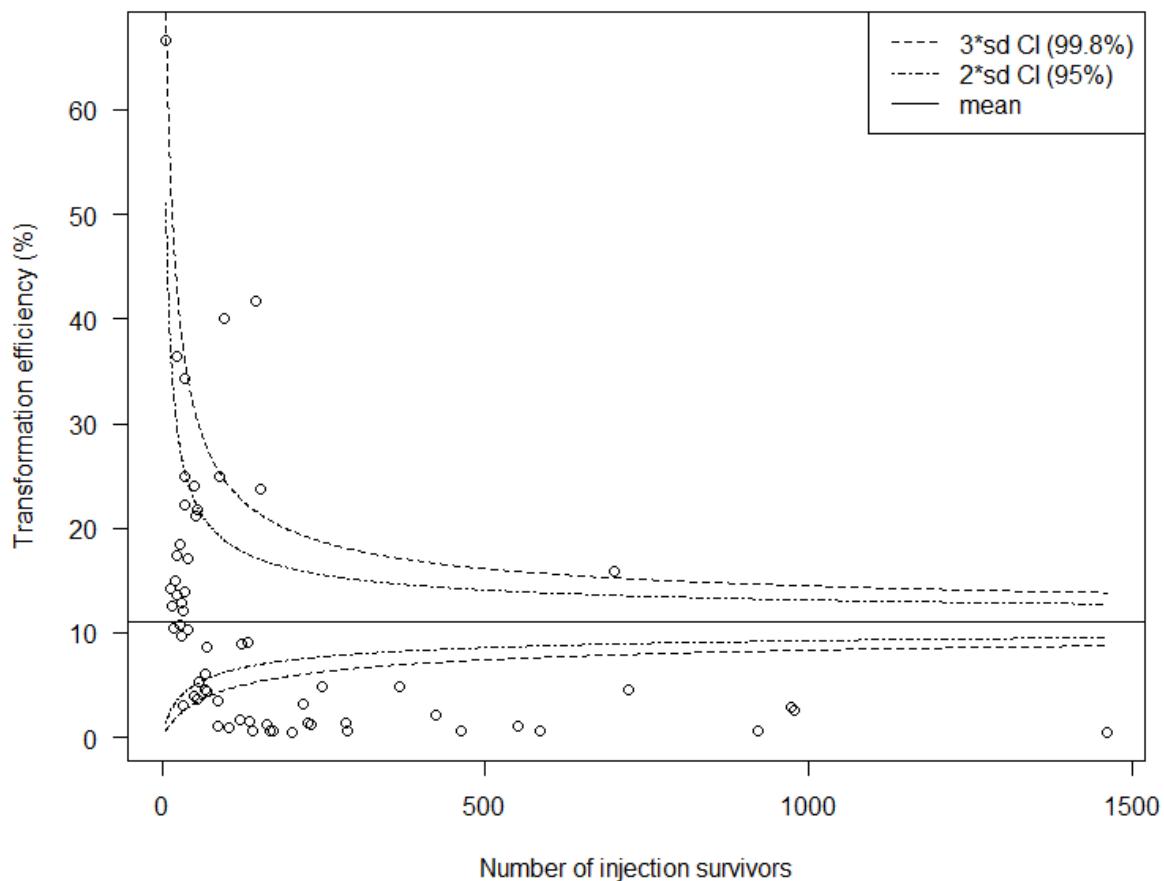


Figure 185. Funnel plot suggesting bias for transformation efficiency data with confidence interval based on sample size generated using Wilson's method (mean – solid line; 95% confidence interval – dotted and dashed line; 99% confidence interval; dashed line). The plot shows all the non-zero transformation efficiency data collected from the literature (64 non-zero experiments plotted with nine zero experiments; another 13 had missing data).

To reveal the inadequacy of this approach we can compare the suitability of the funnel plot with slightly less skewed data, in the survival data for each experiment. However, this also suffers from our treating each species' ability to survive microinjection as homogenous.

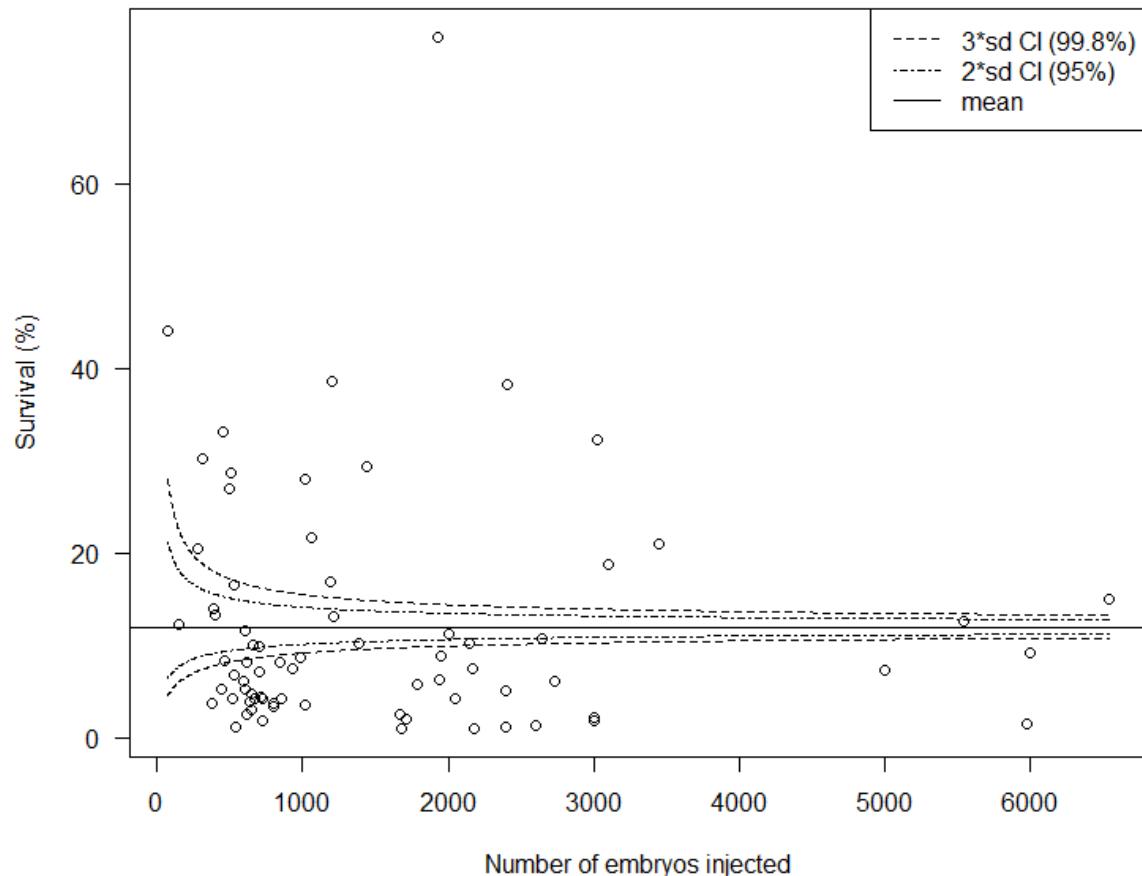


Figure 186. The funnel plot reveals some patient researchers with over 6000 injections in some species before success. Funnel plot of survival data with confidence interval based on sample size generated using Wilson's method (mean – solid line; 95% confidence interval – dotted and dashed line; 99% confidence interval; dashed line). The plot shows all the non-zero transformation efficiency data collected from the literature (64 non-zero experiments plotted with nine zero experiments; another 13 had missing data).

It would be interesting to apply this approach to a data set that was more complete for a given species and transgenesis marker combination as this would control for inter-species variation. To mitigate these issues one can transform each transformation efficiency data and use the length of the 95% confidence interval calculated for the associated experiment and plot them.

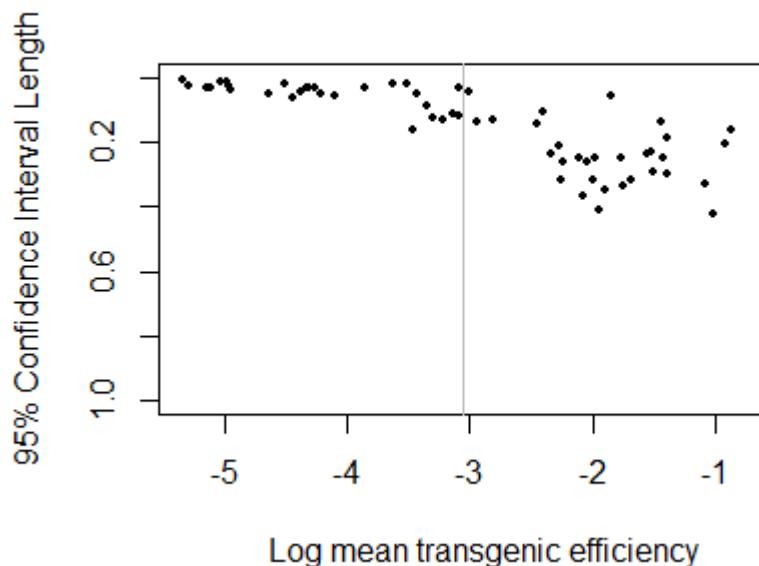


Figure 187. In the absence of publication bias we might expect a symmetrical funnel plot. The bounded nature of the proportion data limits the effectiveness of the plot as the Pearson-Klopper confidence intervals are bounded at zero, limiting the size of the 95% confidence interval which is calculated by subtracting the lower interval from the upper interval at the 95% confidence interval. As the transformation efficiency is not relative to any conventional control, unlike medicine, this removes the relative nature and expected symmetry of the plot.

Figure 187 does not resemble an inverted funnel. This may indicate publication bias; however there are other factors that can lead to an asymmetrical plot. The bounded nature of the proportion data limits the effectiveness of the plot as the Pearson-Klopper confidence intervals are bounded at zero, limiting the size of the 95% confidence interval which is calculated by subtracting the lower interval from the upper interval at the 95% confidence interval. This is borne out by the trend in the figure: as the sample mean efficiency moves away from zero the maximum possible interval size increases. The pooling of species and heterogeneity between research groups may also contribute to the asymmetrical funnel plot. Conventional funnel plots in medicine use a statistic that summarises the treatment effect relative to the control. In this situation no obvious control exists rendering the funnel plot less useful in detecting publication bias.

This bias was not explicitly controlled for in the analysis by statistical corrections. However, the dichotomy between zero and non-zero transformation efficiencies was mentioned and analyses proceeded following sub-setting the data by this criterion.

9.7.1.1.4 QUANTIFYING BIAS

The under-reporting of failure to generate transformants could be quantified by comparing a representative sample of all the experiments that have been conducted, and then comparing the differences to the smaller pool that have been published, noting any differences. Comparing Oxitec grey data with publications; the frequency of failures to transform in Oxitec laboratories relative to other published failures to transform were compared.

A time-line of publications by Oxitec researchers is given. Included are details of the number of unique construct and species combinations whether any zero transformation efficiency experiments were included, the name of the construct injected, the publication details and whether precise injection data was given. References are ordered chronologically and alphabetically. Publications which re-used a previously published construct species combination were omitted from the list. The reference list was derived from the publication record of Luke Alphey, and is assumed to be complete following literature search corroboration.

Table 58. A timeline of Oxitec publications involving transformation events of an insect species using *piggyBac*. Injection data included number of microinjections, number of injection survivors and number of independent transgenic lines including details of construct and helper concentrations used.

| Year | First Author | Species | Unique constructs | Injection data status |
|------|--------------|---------------------------------|-------------------|-----------------------|
| 2005 | Gong | <i>Ceratitis capitata</i> | 2 | absent |
| 2006 | Dafa'alla | <i>Ceratitis capitata</i> | 3 | provided |
| 2006 | Nimmo | <i>Aedes aegypti</i> | 1 | provided |
| 2007 | Condon | <i>Anastrepha ludens</i> | 2 | provided |
| 2007 | Fu | <i>Ceratitis capitata</i> | 2 | absent |
| 2007 | Phuc | <i>Aedes aegypti</i> | 1 | provided |
| 2009 | Morrison | <i>Ceratitis capitata</i> | 1 | absent |
| 2010 | Fu | <i>Aedes aegypti</i> | 2 | absent |
| 2010 | Labbé | <i>Aedes albopictus</i> | 1 | provided |
| 2011 | Simmons | <i>Pectinophora gossypiella</i> | 1 | absent |
| 2012 | Ant | <i>Bactrocera oleae</i> | 1 | incomplete |
| 2012 | Labbe | <i>Aedes albopictus</i> | 1 | provided |
| 2012 | Martins | <i>Plutella xylostella</i> | 4 | provided |

| | | | | |
|------|----------|---------------------------------|---|--------|
| 2012 | Morrison | <i>Pectinophora gossypiella</i> | 3 | absent |
| 2013 | Jin | <i>Plutella xylostella</i> | 2 | absent |

The information above highlights how only non-zero transformation efficiency experiments are likely to be published. The total number of experiments performed at Oxitec permits a true value of the proportion of experiments that result in non-transformation to be compared to an estimate based on the published literature.

Complete information is provided by the columns on the left and compared to the incomplete picture provided by those published experiments.

Table 59. Only Oxitec in-house data is summarised (experiments carried out in partnership with Oxitec are not included). An experiment is defined as a unique construct injected into the given insect species. On occasion an experiment will be unsuccessful in that transgenesis is not achieved, described as a zero transgenics experiment. The number of injection survivors crossed and their progeny screened for transgenics varied.

| Species | Number of zero transgenics (all experiments published) | Total experiments (all experiments published) | Number of zero transgenics (published by Oxitec) | Total experiments (published by Oxitec) |
|---------------------------------|--|---|--|---|
| <i>Aedes albopictus</i> | 0 | 10 | 0 | 2 |
| <i>Aedes aegypti</i> | 5 | 39 | 0 | 4 |
| <i>Plutella xylostella</i> | 9 | 31 | 0 | 6 |
| <i>Ceratitis capitata</i> | 1 | 26 | 0 | 5 |
| <i>Bactrocera oleae</i> | 0 | 5 | 0 | 1 |
| <i>Pectinophora gossypiella</i> | 9 | 35 | 0 | 4 |

The published data underestimates the probability of a zero transformation efficiency in three of the five species (no zeroes in those species least injected or most recently experimented on, *Aedes albopictus* and *Bactrocera oleae*). In the Lepidoptera 27% (18/66) of constructs injected as unique experiments failed to produce any transgenic insects. Furthermore almost half of the publications failed to provide detailed injection data to estimate the transformation efficiency of *piggyBac* with that particular species construct combination (7/15). Unsurprisingly the number of total experiments published by Oxitec is much smaller than the number of experiments carried out. The naïve researcher examining the literature may misjudge the probability of success and the number of intermediary constructs required to achieve a final 'product' line, or

a line that is perceived to be original enough to warrant publication. Further complications may arise due publication being withheld due to patent concerns and intellectual property rights.

Let us assume the position of a naïve researcher interested in creating a diamondback moth transgenic to elucidate if the subset of the data we are provided with is problematic (published literature). Given no zero transformation efficiencies in six experiments, the mean risk of zero transformation efficiency is zero, but what would the 95% upper confidence level (UCL) be? Using a standard frequentist technique (Ludbrook & Lew, 2009) we estimate the UCL to be 39% (Wilson method, R function binom.confint of the package binom, author Sundar Dorai-Raj) (LCL of zero). This compares to the empirical probability of a zero transformation efficiency at Oxitec of 29% (9/31).

If we compare totals then standard research at Oxitec has produced zero transformation efficiency constructs 16% of the time (24/146, or 11-23%, 95% Wilson Confidence Interval) yet Oxitec published research never has zero transformation efficiencies. This is compounded by zero transformation efficiencies rarely being reported by other researchers. This could be mitigated by researchers / editors including / requesting all injection data associated with the development of the genetic construct. Information pertaining to which constructs did not work may be useful in modelling predictors of success or failure when developing new constructs (i.e. do larger constructs have a lower transformation efficiency?).

Published constructs are typically developed piecemeal by incremental increases in knowledge by genetic modules or cassettes that can link together to produce a final construct with a desired phenotype. Although this development is interesting in itself, if a construct does not work it is uninformative as to why it does not work and can be difficult to examine the cause of the failure. This may explain the lack of zero transformation efficiency constructs in the literature, purely because they are perceived to be uninteresting and uninformative.

The aim of this section was to draw attention to the publication bias that exists and to attempt to quantify it for researchers interested in using the *piggyBac* vector for their own insect transgenesis research. The reasons for this are likely to be complex and many.

9.7.1.2 STATISTICAL DESCRIPTION OF THE TRANSFORMATION EFFICIENCY DATA

9.7.1.2.1 APPROPRIATE SUMMARY OF THE TRANSFORMATION EFFICIENCY DATA

A popular summary statistic of the central location of a distribution is the mean, which provides an unbiased estimate (Jones *et al.*, 2009). However, it does suffer from some problems when assumptions of symmetry and normality in the distribution break down. To demonstrate why this is inappropriate a histogram of the data was plotted with an overlay of the density and cumulative density distribution curves (Figure 188). The histogram bin width and number of measurements in total were used to normalise the density curves to the correct height.

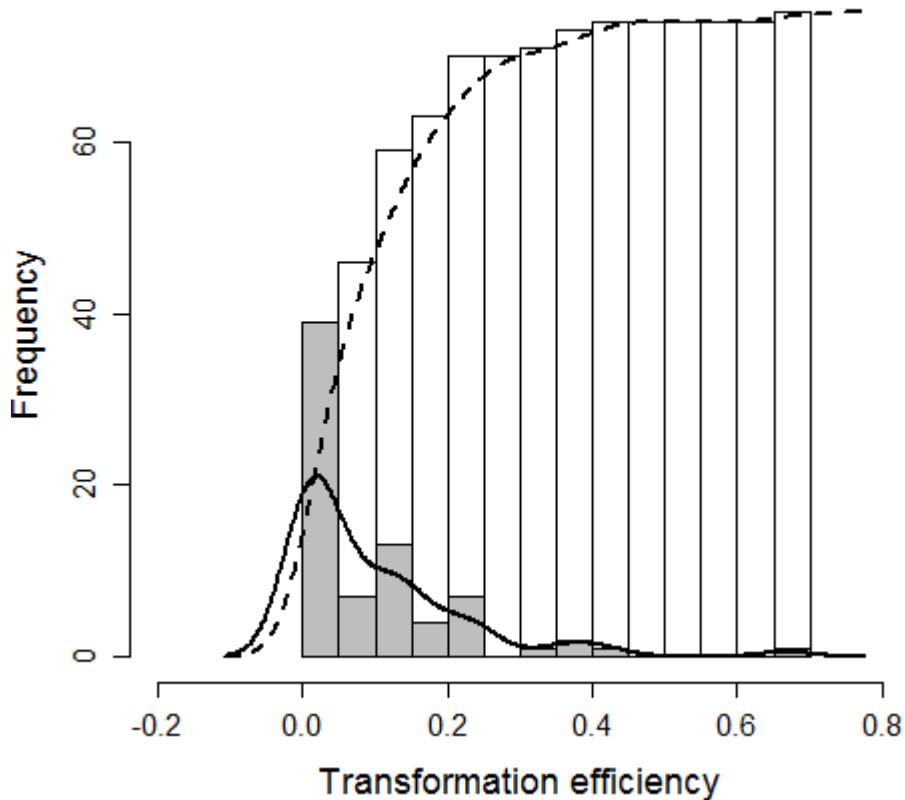


Figure 188. Histogram of the transformation efficiency of the meta-analysis pooled data with bin width of 0.05. The data is positively skewed and bounded between zero and one. A conventional histogram with associated density curve is shaded grey and a solid line. The area under the curve and between the axes integrates to unity and provides a visual representation of the probability of a transformation efficiency falling in a given interval. The cumulative density histogram is also provided as white bars and a dashed density curve. The graph shows why the use of a mean (0.097 and standard deviation of 0.120) to summarise the data is inappropriate as it is asymmetrical. The bounded nature of the transformation efficiency (between zero and one) also creates problems for the normal approximation.

A popular summary statistic of the central location of a distribution is the mean, which provides an unbiased estimate (Jones *et al.*, 2009). However, it does suffer from some problems when assumptions of symmetry and normality in the distribution break down. To demonstrate why this is inappropriate a histogram of the data was plotted with an overlay of the normal distribution curve (Figure 188). The histogram bin width and number of measurements in total were used to normalise the curve to the correct height.

The bounded nature of the data and the heavy skew means the normal distribution is a very poor model of the data. The skew and leptokurtosis render the mean useless as a descriptive statistic of the central location of the transformation efficiency distribution in insects injected with *piggyBac* vector.

Assuming history can be used to predict the future (ignoring the problem of induction) the expected value of the transformation efficiency of a random construct can be described. For the

moment we ignore any effects of differences between the experimental conditions from which the transformation efficiencies were derived and assume that transformation efficiency of a *piggyBac* vector is described by a mean efficiency and a random error term.

The data can be modelled discretely by using intervals of transformation efficiency. The table below gives the probability mass function for transformation efficiency of constructs published in the literature (Table 60). It summarises the probability $P(y)$ of achieving a transformation efficiency in a given interval, the discrete random variable (y).

Table 60. The probability mass function of the transformation efficiency data found in the literature offers a complete empirical probability mass function version in graphical form. Technically zero is a point not an interval. The probability (0-1) that a transformation efficiency of a publication randomly sampled from the literature sample will fall into given intervals (3.s.f.).

| Transformation efficiency interval (y) | 0 | >0-0.1 | >0.1-0.2 | >0.2-0.3 | >0.3 |
|--|-------|--------|----------|----------|-------|
| $P(y)$ | 0.135 | 0.486 | 0.230 | 0.095 | 0.068 |

The empirical cumulative distribution function (ECDF) can be calculated and plotted for the transformation efficiency data (to convert to percentages multiply by one hundred). It provides all of the information from the data in a graphical format which can be easily read from. For example if a researcher wants to estimate the probability that a transformation efficiency greater than 0.2 will be achieved, one would trace up from $p=0.2$ and then along to the associated $f(p)$ value of 0.8 (it can also be computed given the ECDF; a re-mapping of the data). This tells us that approximately 0.8 of all observations are below a transformation efficiency of 0.2, therefore the fraction of observations greater than 0.2 transformation efficiency is only $(1-0.8)$ approximately 0.2 of the observations.

Transformation efficiency is a continuous variable; accordingly it would be preferable to fit a probability density function with a suitable mathematical model to describe the data; compressing the data into a function of two numbers. The model could then be used to describe probabilities with given values of the transformation efficiency by integration of the curve.

It can be difficult to arrive at a decision as to what distribution best characterises the data, as normally one would make a decision based upon the data distribution using a histogram. The problem with the histogram is that the interval size of the bins can affect the geometry of the distribution. An alternative approach is to consider how the transformation efficiency value is generated and its mathematical properties, and consult the literature for strategies to model a fit to the data.

Conventionally for proportion data the logit transformation was used to normalise the distribution of proportion data. Different probabilities of success, p , are transformed into logit ($\ln(p)$) (or $\ln(p/(1-p))$). As there are a large number of zeroes, this may not be appropriate as the logit of zero is minus infinite, adding one to all data causes additional problems (O'Hara and Kotze, 2010). The logit transformation can also make interpretation difficult so an alternative strategy was preferred.

9.7.1.2.2 PROBABILITY DENSITY FUNCTION OF THE TRANSFORMATION EFFICIENCY DATA

Transformation efficiency is a derived variable. It is assumed to be emergent from a series of independent and identically distributed random variables ($X_1, X_2, X_3 \dots X_n$; range: success (transgenic) = 1, failure (non-transgenic) = 0), with constant probability of success p .

Transformation efficiencies of zero or one were removed from the data and a subset created of transformation efficiencies with “no zeroes”. This subset provides us with a strategy to generate a probability density function of the probability of obtaining a transformation efficiency value on condition that transgenesis was achieved (transformation efficiency is non-zero, $p > 0$ and < 1). According to Table 60, the probability of achieving a non-zero transformation efficiency is 0.865 ($1 - 0.135$).

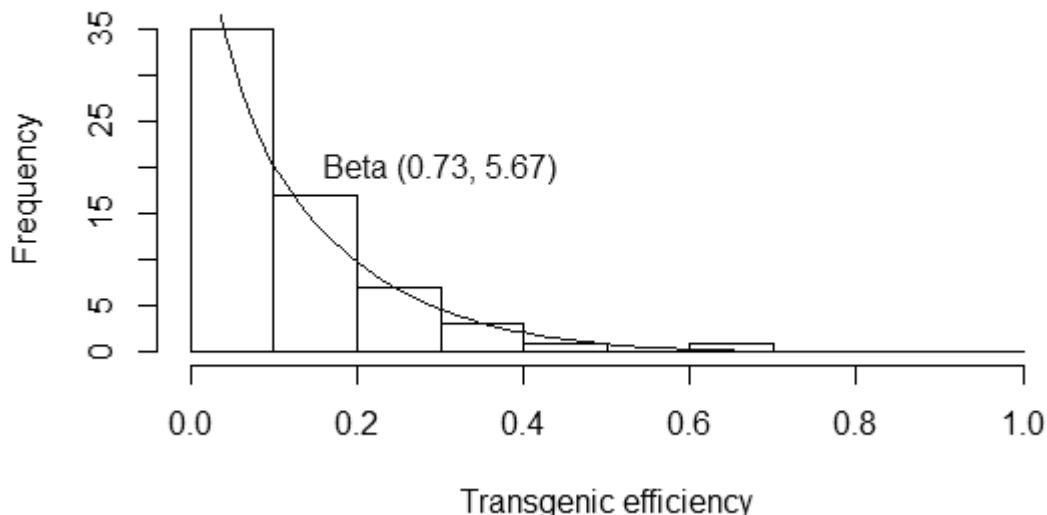


Figure 189. A histogram of the non-zero transformation efficiencies of experiments from the systematic literature review. A beta distribution curve with parameter estimates was overlaid, fitted from the data using “fitdistr” function from the MASS package in R (Brian Ripley, 1998). The bin width and number of measurements in total forming the histogram were multiplied by the beta distribution so that the curve could be normalised to the correct height. The area under the curve and between the axes integrates to unity and provides a visual representation of the probability of a transformation efficiency falling in a given interval.

This distribution function can be used to estimate the probability of achieving transformation efficiencies in any specified interval using the “pbeta” function (Catherine Loader) in R with the shape 1 and shape 2 parameters for the beta distribution derived from the data using maximum likelihood methods (shown in blue in Figure 189). It is a more convenient and succinct version of the ECDF, as it summarises the information into a function of two numbers. This information is useful for decision-making and informing researchers as to the likely number of fertile injection survivor crosses required for success. This number can then be multiplied by the injection-related mortality in the species of interest to determine an appropriate number of

injections required for successful transformation (given the data or prior information in literature). This may facilitate estimating research costs. The beta distribution format of the probability density function is amenable to use in Bayesian analyses and decision-making, increasing the utility of these findings.

Going back to Fraser (2012); using this model the probability of the non-zero transformation efficiency of a *piggyBac* vector insertion lying between 0.001 and 0.1 is 0.586; most efficiencies are in this interval when the data is summarised using a beta distribution.

To summarise - if a researcher is working with a species for which no literature is available to suggest a transformation efficiency estimate, then Figure 189 provides a distribution of the expected transformation efficiency given past successes in insects, dependent that the species is amenable to transformation and the construct can indicate successful integration (and the construct is non-lethal).

To validate the model two approaches are feasible.

- The dataset could be randomly split in half with one subset used to build the model and the second subset to test it. The leave-one out cross validation method could also be implemented to assess model validity.
- An additional dataset could be accrued over time and tested against the model shown here.

9.7.1.3 BAYESIAN ANALYSIS OF TRANSFORMATION EFFICIENCY – A TOY EXAMPLE WITH *TRIBOLIUM CASTANEUM*

9.7.1.3.1 TRIBOLIUM CASTANEUM TRANSFORMATION EFFICIENCY ESTIMATE FROM SEVERAL EXPERIMENTS

Although the variation between experiments is large it seems a plausible range of likely values of the transformation efficiency can be observed for a species which has seen numerous independent experimental instances of transformation particularly if the same promoter-fluorescent protein combination were used to report success. The *Tribolium castaneum* transformation efficiency lies between 0.2-0.4 making it amenable to Bayesian analysis. Of course this subjective interpretation should be revised and objective methods based on meta-analysis strategies of combining summary statistics from different experiments should be employed. This information can be used to produce a prior transformation efficiency estimate or preferably a probability distribution of likely transformation efficiencies for a particular species given enough experiments. This could then be used as a decision tool for estimating the number of G_0 required for likely success in *Tribolium castaneum*.

The intended outcome is an appropriate probability density function for use in Bayesian analysis. An approximate answer to the right question is desired, over a precise answer to the wrong question (the frequentist approach would not permit prediction of a parameter associated with an event that has not yet happened, O'Hagan & Luce, 2003). A controversial assumption includes that differences in the design of the constructs have negligible impact on transformation efficiency (based on the same transgenesis module being used within each

species). The data is derived from the same lab so methodologies are broadly similar with some discrepancies in plasmid and helper concentrations, for example.

9.7.1.3.2 BAYESIAN ANALYSIS FOR A BERNOULLI PROCESS

An example of how to merge transformation efficiency data for a single species from different experiments will be demonstrated using a Bayesian approach (O'Hagan, 2009) following methods of Winkler *et al.*, 2002. The use of a prior distribution and the Bayesian approach is robust even in the face of the zero-numerator problem (Ludbrook & Mew, 2009), providing excellent coverage in situations likely to occur in insect transgenesis research.

For consistency, the methods will be described using the symbols in Winkler *et al.*, (2002). For a more detailed discussion of the technique, the reader is referred the details in the paper (the original approximation for the binomial process using the beta distribution used in Winkler is after Ishii and Hayakawa (1960)).

Imagine you were in the position of Lorenzen *et al.*, in 2003. In an earlier experiment they had successfully transformed *Tribolium castaneum* using *piggyBac*. They generated 38 transgenic lines with 95 G₀. Going into the experiment they have an estimate of the transformation efficiency in *Tribolium castaneum* based on the maximum likelihood estimate of the mean. The Bayesian approach requires a prior distribution which can be estimated conveniently:

$$f(p) = p^{a-1}(1-p)^{b-1}/B(a,b) \text{ for } 0 \leq p \leq 1$$

The estimate of the mean is $a / (a+b)$ with a the number of successes and b the number of failures. The true mean, p , is the probability of success bounded between 0 and 1.

The information about p represented by the prior Beta(a, b) distribution can be interpreted as the equivalent of having seen a transgenics given $a + b$ fertile injection survivors crossed.

Given the prior, in a later experiment, if we observe r independent transgenic lines given n trials (the data is the likelihood) the posterior distribution is Beta($a + r, b + n - r$).

Prior: after Lorenzen (2003), $a = 38$, $b = 95$.

The prior was chosen based on the only relevant empirical data available at the time.

Likelihood: after Lorenzen (2003), $r = 36$, $n = 152$.

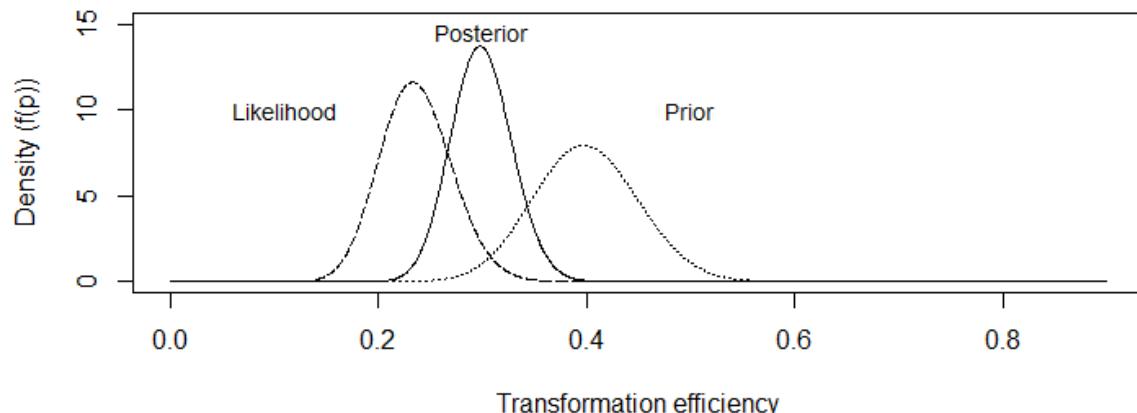


Figure 190. Bayesian tri-plot for the mean transformation efficiency probability density in *Tribolium castaneum*. The prior was formulated using 38 transgenics given 95 fertile G₀ crosses (after Lorenzen *et al.*, 2003), the likelihood represents the data of 36 transgenics given 152 fertile G₀ crosses (after Lorenzen *et al.*, 2003). The prior and likelihood are combined using Bayes' Theorem to create the Posterior distribution which provides a probabilistic parameter estimate of the transformation efficiency in *Tribolium castaneum* given previous information and recent experimental evidence. R code: tc_g0_lambda.R

Bayes' theorem allows us to combine information from two different sources using explicit methodology. The resulting posterior distribution permits inferences about the parameter of interest; the transformation efficiency of *piggyBac* in *Tribolium castaneum*. Observational studies can provide strong inference for parameter estimation but are limited in assessing causality.

It follows that the transformation efficiency is described by the posterior probability distribution with the mode at 0.298. This provides an intuitive probability distribution (in contrast to a confidence interval, see Morey *et al.*, 2015) providing plausible parameter estimates with the integral of the distributions. The reliability of the estimate is based on the similarity between the experiments and methodology, as well as assumptions of the binomial model of insect transgenesis success described above. Due to research institution-level heterogeneity it is preferred to use data from the researcher's own institution when generating a predicted transformation efficiency.

9.7.1.4 OXITEC DATA

9.7.1.4.1 STRUCTURE

The term data set (or dataframe) defines a set of measurements taken from some environment or process. The data set will take the form of the collection of measurements on n objects, comprising the rows (unique constructs injected into a unique species). The columns are made up of p variables which represent the columns and variables of the data set; which has size, n times p (after Little & Rubin, 2002). The data was represented as a dataframe in R, with missing values assigned "NA", for not available, missing values were due to a failure to observe or record

(Adler, 2010). The data is comprised of empirical distributions of variables, based on a finite number of empirical observations that will continue to expand and update through time. A well structured data frame lends itself to continual updating facilitating future analysis and hypothesis generation by interested researchers.

9.7.1.4.2 DEALING WITH MISSING VALUES

In the data set some of the entries were not observed (constructs could be abandoned before success rates were measured or entries were not made and or not recorded, etc.). These values are missing in the sense that there were actual values that could have been assigned to the entry if observation or recording techniques were not lacking. These entries were replaced with a suitable place holder given the statistical software used for analysis (NA when using R; Adler, 2010). After Little & Rubin (2002) it is assumed that “missingness indicators hide true values that are meaningful for analysis”.

One simple way around the problem of missing values is to exclude them from the analysis. This strategy is termed “complete-case analysis”. Although straightforward this approach can be problematic given a small data set or the nature of the missingness, known as the missing data mechanism. If the missing data mechanism has some dependence on the values of the data the mechanism is called “not missing at random” (NMAR). The missing data from our data-set seems to be NMAR. As suggested earlier this is due to the nature of the research environment; if a very similar construct has been very successful in generating transgenics (or in generating transgenics that do not have the predicted or desired phenotype) observations were not made. Data were transliterated chronologically, assuming they were entered originally in a chronological order, based on this assumption the missingness of data seems to be clustered around certain time periods.

This NMAR makes it difficult to estimate missing values. Accordingly complete-case analysis was preferred. Following removal of all rows any NA or NaN (not a number, for example when dividing by zero) value in any column 149 complete entries or rows of the 166 remained each described by the five variables (species, construct ID, injections, injection survivors and number of lines generated) of interest for this study (27 variables were included for posterity). Due to these differences, systematic error between missing and non-missing data may occur, confounding analysis. The probability that an observation is missing is likely to be dependent on the outcome being studied.

This was too conservative; ideally any entry would be removed if it tested NA given the specific analysis being carried out (i.e. for a particular explanatory variable involved in the analysis). This approach, called pairwise deletion, was preferred.

9.7.1.4.3 THE EXPERIMENTAL UNIT AND THE EVALUATION UNIT

The experimental procedure for generating transgenics is summarised in Chapter 2, and was kept constant except for species type of embryo and the construct injected. The species of the embryo forced experimental constraints on variables that could be controlled including rearing practices and ratios at which the injection survivors were pooled with wild type insects to give rise to G₁. The nomenclature used is based on suggestions of Hurlbert (2009):

an experimental unit is defined as “The smallest system or unit of experimental material to which a single treatment … is assigned … and which is dealt with independently.”

The construct label itself is a random factor. The size of the construct injected and presence or absence of certain genes were also considered during analyses. Each trial (or experimental unit) (n) is summarised in the master data frame (<https://github.com/mammykins/piggyBac-data>). Each row or observation in the data set comprised a large number of independent repeats of “n” (minimum = 100, median = 1480, maximum = 12000) or injections. The data set is comprised of almost 297,755 experimental units which provide numerical, random categorical, fixed categorical and derived variables on 166 observations.

9.7.1.4.4 THE NUMBER OF MICROINJECTIONS AND METHODOLOGICAL CONSIDERATIONS

The Diptera have received 161,132 embryo microinjections between them. This is split between two teams or departments at Oxitec: one specialising in transformation of fruit flies, the other in mosquitoes. Due to challenges in embryo collection and preparation compared to the Lepidoptera, the injections tend to be conducted by teams of people rather than individual researchers, with regular swapping of roles to maintain motivation and concentration throughout (the moth transformation team carried out embryo collection and injection as individuals). It is unclear how random this process was and how much variation a researcher carrying out the injections has on this data exploration. Methodologies were similar throughout the collection of the data as all participants sought process optimisation by careful control of variables and monitoring or recording of any changes to methods. Importantly the marker modules (promoter and fluorescent protein to indicate successful germline transformation) were relatively constant within a species.

Almost 131,400 lepidopteran (pink bollworm, diamondback moth and *Tuta absoluta*) embryo injections made up the Lepidoptera data sub-set. Of these, 211 produced unique insertion events that were detected over the 6 years.

The Coleoptera despite being the most speciose order are relatively neglected with only 5,000 embryo injections in *Tribolium castaneum*.

9.7.1.4.5 BIAS CONSIDERATIONS WHEN COMPARING THE META-ANALYSIS AND OXITEC DATA

Cross referencing the transformation efficiency data against the number of injection survivors used showed a correlation between the highest transformation efficiencies and lowest number of injection survivors; the experiments with the lowest number of trials produced higher efficiencies, as shown in the funnel plots (Figure 191). This provides an important caveat to consider when inferring a trend from this data.

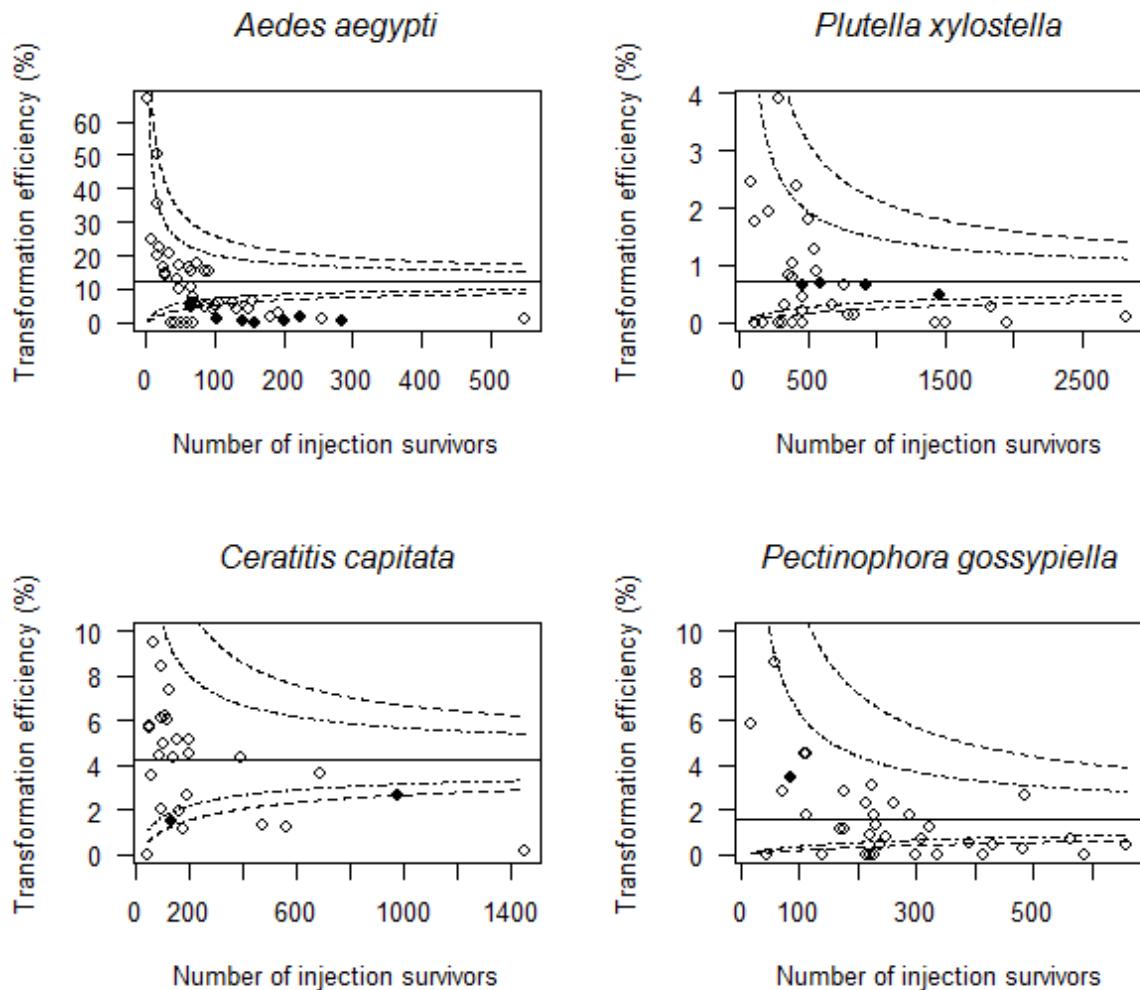


Figure 191. The mean and binomial confidence intervals appear to overestimate the transformation efficiency consistently in all four well-studied species shown (*Aedes aegypti* = 30, *Plutella xylostella* = 33, *Ceratitis capitata* = 26, *Pectinophora gossypiella* = 35). The axes are not constant so take care when comparing between species. The data includes some experiments included in publications highlighted black. The funnel plot provides transformation efficiency data with confidence interval based on sample size generated using Wilson's method (mean – solid line; 95% confidence interval – dotted and dashed line; 99% confidence interval; dashed line). A benefit to the funnel plot is that it highlights the ever present danger of mistaking variation due to chance for correlation or causation.

Some of the Oxitec experiments have been published with the summary statistics included in the meta-analysis data. There published experiments are compared to all other experimental data for that species transformation efficiency at Oxitec (Figure 191). The skew also reinforces that the mean is a poor measure of central tendency for transformation efficiency with the median preferred.

9.7.1.4.6 SURVIVAL DATA CONSIDERATIONS

As the median of each distribution approaches 0.5 the variance appears to increase while the skew decreases, suggesting a normal approximation could be viable for the distribution of a samples survival in *Plutella xylostella* and *Pectinophora gossypiella*. Conversely those species distributions with a nearer zero median may be better represented by a non-normal distribution.

Three of the species have outliers – constructs injected that were more than 1.5 times the inter-quartile range (IQR) away from the median. Outliers occur in those species with a greater number of data; as they are simultaneously improbable and extreme. For those species with smaller IQR it is easier to predict a future expected survival following microinjection with greater accuracy. Those four species with the lowest IQR were also the four species with the lowest number of constructs injected and should be excluded from the previous assertion.

Walton *et al.*, 1987 showed a relationship between the diameter of the needle and the probability of embryo lysis or cellular disintegration although this was not consistent between species. It was speculated that this was due to the relative size of the embryo to the needle and the taper of the needle. The different species at Oxitec are injected with needles made from different programme specifications on a glass-capillary needle puller (programmes have changed through time). Despite this confounding variable, it would be interesting to compare species embryo average dimensions to needle diameter ratio with embryo survival.

Recent additions to the Oxitec research portfolio have fewer experiments and lower variation (specifically *Tuta absoluta*, *Tribolium castaneum*, *Bactrocera oleae* and *Drosophila suzukii*). Those with more injections tend to have greater variation. This could reflect improved injection and husbandry systems for working with the particular insect species through time (or the natural tendency for the range of a rare event statistic to increase through time). Interestingly the mosquitoes have been injected for many years but still show a low survival of typically less than 10% this suggests there may be a ceiling to survival for microinjection.

9.7.1.4.7 CONSIDERATIONS FOR THE DECISION MODEL

The more variable the survival the less useful it is to ask – given n injections how many transgenic lines can one expect? For these species it may be preferred to ask – given a number of G_0 , how many transgenic lines can one expect? This was considered an important requisite of a decision model. Accordingly the survival can be set to one to assess the transformation efficiency given a number of G_0 .

It is also important to note that the decision model was constructed in a way to make it useable to other researchers (they provide fixed input parameters based on their empirical evidence or information in this review) and should not be limited by potential lab-specific biases that relying on only the Oxitec data could create. For species for which many prior experiments exist it may be preferred if the survival and transformation efficiency of each simulated embryo experiences a parameter value drawn from the empirical distribution weighted by the size of the denominator for that species. This approach was investigated but sacrificed speed for precision.

9.7.1.4.8 ASSESSING DECISION MAKING USING INCOMPLETE INFORMATION

The model was run for 500 injections and 1000 (top and bottom rows of Figure 192) injections for both the published data and Oxitec data parameters (left and right columns).

Table 61. The diamondback moth parameters to be used in two separate uses of the model to compare differences or the bias produced by reliance on an incomplete data set (published). Both survival and transformation efficiency medians are given for the published and Oxitec datasets.

| | Published | Oxitec |
|----------------------------------|------------------|---------------|
| Survival (s) | 0.3839 | 0.2800 |
| Transformation efficiency | 0.0065 | 0.0043 |

The parameters derived from the literature reveal the survival estimate lies outside the inter-quartile range of our complete survival data set. The transformation efficiency estimate compares more favourably and is close to the median derived from the complete data set (Figure 192).

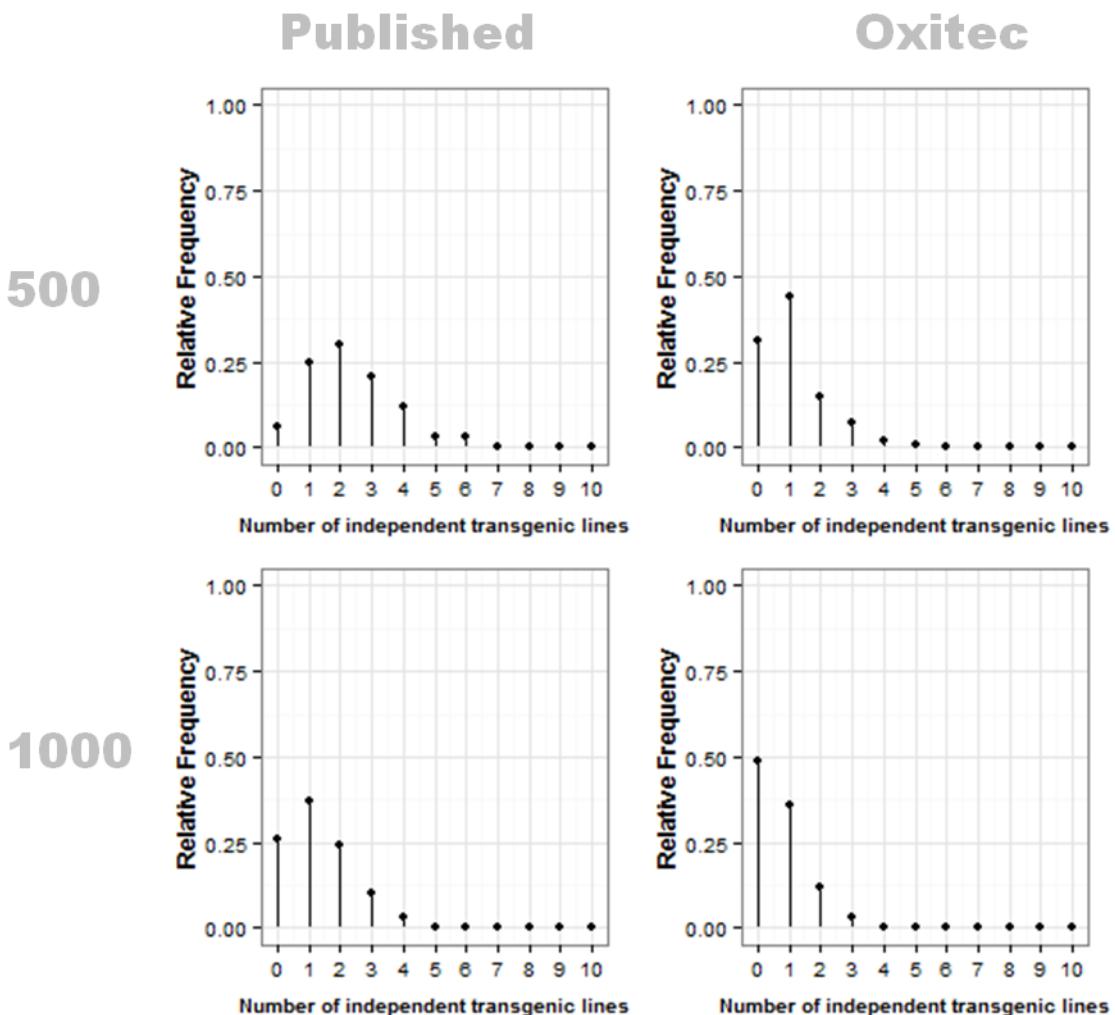


Figure 192. Output from the goldilocks simulation model – helping researchers to get the number of injections just right. The top and bottom row are simulations involving 500-1000 injections, respectively. The left column uses the median statistics from the published data and the right from a more complete Oxitec data set identifying publication bias. The literature provides an overly optimistic view of the chances of successful transformation given a number of injections.

The relative frequency is the number of simulations that resulted in the specified number of transgenic lines divided by the total number of simulations. The model demonstrates how basing ones decisions on published data could lead to a failure to transform due to insufficient injections. The simulations contrast with 500 injections whereby the published data suggests a 25% chance of failure compared to a 50% chance using the complete data set. Doubling the workload to a thousand injections almost halves the probability of failure for both the published and complete data set parameters.

This approach applied to other well-documented species at Oxitec revealed the literature overestimates the transformation efficiency for the Lepidoptera (*Plutella xylostella* and *Pectinophora gossypiella*) and underestimates for *Aedes aegypti* and *Ceratitis capitata*. The

moral is to use as complete a dataset as possible to ensure a better understanding of what is typical.

9.7.1.4.9 APPLYING THE MODEL

Download R, R studio, the Shiny and ggplot2 packages in R and the Goldilocks application. Run the application. Carry out the number of injections necessary to reduce the risk of zero transgenic lines to an acceptable relative frequency. This model should be used as a heuristic and should not replace thinking about the specifics of the relevant problem a researcher faces. Alternatively an online version is available at <https://mammykins.shinyapps.io/App-gold>.

The model has many uses and can be used in similar situations for other vectors or even other modes of transformation, such as the gene gun. If summary statistics are available for the survival and transformation efficiency, then this can be co-opted for other methods and facilitate decision making for gene insertion approaches.

This model can also be used to provide a p value for justifying when one has carried out sufficient injections for a particular construct.

9.7.1.4.10 A PREFERRED MODEL USING ALL THE DATA AT THE EXPENSE OF TIME TO SIMULATE

The previous model used two parameter estimates to drive the model; the median survival and the median transformation efficiency. Given many experimental results we can improve precision by using all experimental results. This is achieved by the model drawing from a vector containing all experimental results with the probability of sampling a particular datum weighted by the number of trials used in that particular experiment. This approach was preferred as it used all empirical data available but was not skewed by experimental parameter estimates based on a small number of trials. This provides extra precision at the expense of time to run the simulation. Accordingly for a live web-based application the earlier method was preferred due to a faster computing time. Comparing simulation outputs did not reveal a difference for those insects with many data. For those insects with few data the models are likely to vary in their output (data not shown).

9.8 GLMM DETAILS

The two main fixed effects were the type of injection mix treatment applied to the embryo and the experimenter who carried out the screening of the embryos post-injection (referred to as “screener”) (sometimes both screened the same slide/embryo independently). The random effect of slide was included and assigned an arbitrary intercept of 1. The response variable is the transient status of the injected embryo up to 3 days post-injection (coded as zero if it showed no transient-expression for all three days or coded as one if it showed transience on at least one day). The model was not over-dispersed (residual deviance/residual degrees of freedom < 1) validating the choice of link function.

The maximal model: lmer (transient status ~ treatment + screener + (|slide), family = binomial) was fitted in R.

On inspection of the random effect 'slide', each slide contributed a similar number of instances to the data. However, the mixed effect model including the random effect had a lower AIC score compared to a GLM not including the random effect term. However, AIC depends on counting parameters, comparing AIC values with different random-effects structures can be suspect.

Model simplification was continued by comparing the maximal GLMM with a GLMM with the screener effect removed. No significant difference in explanatory power was lost at the expense of screener removal ($\chi^2=1.32$, $p=0.25$).

Non-significant terms were removed in a step-wise manner. The new model was compared to the previous using ANOVA, if no significant difference in explanatory power occurred the updated model was preferred using Occam's razor. Interactions were removed first and then main effects in a step wise process until removal of a term resulted in a loss in explanatory power (for model simplification see Crawley, 2005).

Following model simplification the minimum adequate model was: lmer (transient status ~ treatment + (|slide), family = binomial).

The residual deviance of the model was 1385 with 1842 degrees of freedom. The screener was unimportant, as in the screener did not significantly affect the likelihood of a treatments transient-expression status. However, this does not mean that the screeners agreed on the status of every embryo that both screened.

9.9 PAPERS

Chapter 2 has been redrafted and sent for review at the journal of Insect Molecular Biology. The ability to publish the other chapters is limited by the sensitive nature of the information but they are likely to contribute to a patent for the paternal effect system. I would like the work to contribute to a larger publication that could contain details of the paternal effect system in insects at Oxitec.

Gregory, M., Alphey, L., Morrison, N. and Shimeld, S. (2015). Insect transformation with piggyBac: getting the number of injections just right. Insect Molecular Biology. Accepted.

9.10 CHAPTER COVER PAGE PHOTOS

Chapter 1: Sunrise over a cabbage field in Lompoc, California – credit: Kevin Cole (www.kevinlcole.com).

Chapter 2: Microinjection of diamondback moth – credit: Adam Walker, Oxitec.

Chapter 3: diamondback moth pupae co-transformed with DsRed, AmCyan and ZsGreen fluorescent protein markers – credit: Adam Walker, Oxitec.

Chapter 4: diamondback moth flying collage – credit: Derric Nimmo, Oxitec.

Chapter 9 | Appendices

Chapter 5: *Tribolium castaneum* – credit: Derric Nimmo, Oxitec.

Chapter 6: *Tctra* dsRNA induces RNAi in *Tribolium* – credit:

Chapter 7: A better world through DNA®– credit: Intrexon, <https://www.dna.com/>

CHAPTER 10: REFERENCES

- Adelman, Z.N., Jasinskiene, N. & James, A. a, 2002. Development and applications of transgenesis in the yellow fever mosquito, *Aedes aegypti*. *Molecular and biochemical parasitology*, 121(1), pp.1–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11985858>.
- Adler, J., 2010. *R in a nutshell* First. M. Loukides, ed., Sebastopol: O'Reilly Media Inc.
- Aito, T.S., Irai, K.H. & Ay, M.O.W., 2005. The rice water weevil , *Lissorhoptrus oryzophilus* Kuschel (Coleoptera : Curculionidae). *Applied Entomology and Zoology*, 40(1), pp.31–39.
- Al-Dosari, M.S. & Gao, X., 2009. Nonviral gene delivery: principle, limitations, and recent progress. *The AAPS journal*, 11(4), pp.671–681.
- Allen, M.L. et al., 2004. piggyBac transformation of the New World screwworm, *Cochliomyia hominivorax*, produces multiple distinct mutant strains. *Medical and veterinary entomology*, 18(1), pp.1–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15009439>.
- Alphey, L. et al., 2013. Genetic control of *Aedes* mosquitoes. *Pathogens and global health*, 107(4), pp.170–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23816508>.
- Alphey, L., 2014. Genetic control of mosquitoes. *Annual review of entomology*, 59, pp.205–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24160434> [Accessed July 17, 2014].
- Alphey, L. et al., 2006. Genetic technologies to Enhance the Sterile Insect technique (SIT). In *Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*. Salvador, pp. 319–326.
- Alphey, L. et al., 2008. Insect population suppression using engineered insects. *Advances in Experimental Medicine and Biology*, 627, pp.93–103.
- Alphey, N., Alphey, L. & Bonsall, M.B., 2011. A model framework to estimate impact and cost of genetics-based sterile insect methods for dengue vector control. *PloS one*, 6(10), p.e25384. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3187769&tool=pmcentrez&rendertype=abstract> [Accessed September 3, 2013].
- Anon, 2005. Fundamentals. In *Statistics and introduction using R*. Chichester, West Sussex: Wiley & Sons, p. 12.
- Ant, T., 2013. Genetic control of the olive fruit fly , *Bactrocera oleae*. The University of Oxford.
- Arthur, F.H., 1996. Grain protectants: Current status and prospects for the future. *Journal of Stored Products Research*, 32(4), pp.293–302.
- Arunkumar, K.P. & Nagaraju, J., 2011. Drosophila intersex orthologue in the silkworm, *Bombyx mori* and related species. *Genetica*, 139(1), pp.141–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21120683> [Accessed October 13, 2011].
- Asokan, R. et al., 2013. Response of various target genes to diet-delivered dsRNA mediated RNA interference in the cotton bollworm, *Helicoverpa armigera*. *Journal of Pest Science*, 87(1), pp.163–172. Available at: <http://link.springer.com/10.1007/s10340-013-0541-7> [Accessed December 2, 2014].

- Atkinson, P.W., 2002. Genetic engineering in insects of agricultural importance. *Insect biochemistry and molecular biology*, 32(10), pp.1237–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12225914>.
- Bagla, P., 2010. Hardy Cotton-Munching Pests Are Latest Blow to GM Crops. *Science*, 327, p.1439.
- Baird, G. & Zacharias, D., 2000. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proceedings of the*, 97(22), pp.11984–9. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC17281/> [Accessed June 27, 2012].
- Balhorn, R., 2007. The protamine family of sperm nuclear proteins. *Genome biology*, 8(9), p.227. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2375014/> [Accessed October 24, 2012].
- Banasiak, U. et al., 2010. Plant protection products and their residues. Aspects of consumer safety in context of the new EU regulations. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 53, pp.567–576.
- Banerji, J., Rusconi, S. & Schaffner, W., 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, 27(2 Pt 1), pp.299–308. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6277502>.
- Barckmann, B. et al., 2013. Three levels of regulation lead to protamine and Mst77F expression in Drosophila. *Developmental Biology*, 377(1), pp.33–45. Available at: <http://dx.doi.org/10.1016/j.ydbio.2013.02.018>.
- Bargielowski, I., Alphey, L. & Koella, J.C., 2011. Cost of Mating and Insemination Capacity of a Genetically Modified Mosquito Aedes aegypti OX513A Compared to Its Wild Type Counterpart. *PloS one*, 6(10), p.e26086. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3219117/> [Accessed October 30, 2011].
- Barrett, L.W., Fletcher, S. & Wilton, S.D., 2012. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences*, 69(21), pp.3613–3634.
- Bates, D. et al., 2015. Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, In press. Available at: <http://arxiv.org/abs/1406.5823>.
- Baum, J. a et al., 2007. Control of coleopteran insect pests through RNA interference. *Nature biotechnology*, 25(11), pp.1322–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17982443> [Accessed November 1, 2012].
- Bazuin, S., Azadi, H. & Witlox, F., 2011. Application of GM Crops in Sub-Saharan Africa: Lessons Learned from Green Revolution. *Biotechnology Advances*, 29(6), pp.908–912. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0734975011001212> [Accessed August 2, 2011].
- Beddington, J., 2010. Food security: contributions from science to a new and greener revolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1537), pp.61–71. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2842707/> [Accessed July 18, 2012].

- Behrens, J.T., 1997. Principles and procedures of exploratory data analysis. *Psychological Methods*, 2(2), pp.131–160. Available at: <http://doi.apa.org/getdoi.cfm?doi=10.1037/1082-989X.2.2.131>.
- Bellen, H.J. et al., 2011. The Drosophila gene disruption project: progress using transposons with distinctive site specificities. *Genetics*, 188(3), pp.731–43. Available at: <http://www.ncbi.nlm.nih.gov/article/fcgi?artid=3176542&tool=pmcentrez&rendertype=abstract> [Accessed March 29, 2013].
- Benson, K. & Hartz, A., 2000. A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine*, 342(25), pp.1878–1886. Available at: <http://www.nejm.org/doi/full/10.1056/NEJM200006223422506> [Accessed April 24, 2013].
- Berger, Z., 2010. Bayesian and frequentist models: legitimate choices for different purposes of clinical research. *Journal of evaluation in clinical practice*, 16(6), pp.1045–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20860579> [Accessed October 30, 2012].
- Berghammer, A.J. et al., 1999. A universal marker for transgenic insects. *Cell*, 402(November), p.370.
- Berghammer, A.J. et al., 2009. Red flour beetle (*Tribolium*) germline transformation and insertional mutagenesis. *Cold Spring Harbor protocols*, 2009(8), p.pdb.prot5259. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20147235> [Accessed August 5, 2011].
- BeVier, G., 2012. Global food systems: feeding the world. *Reproduction in domestic animals = Zuchthygiene*, 47 Suppl 4, pp.77–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22827353> [Accessed October 5, 2012].
- Bevis, B.J. & Glick, B.S., 2002. Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nature biotechnology*, 20(1), pp.83–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11753367>.
- Bilha, J., Brancalhão, R. & Conte, H., 2012. Testicles fusion of *Diatraea saccharalis* F.(Lepidoptera; Crambidae) during post-embryonic development. *Journal of the Entomological ...*, 14(3), pp.1–7. Available at: [http://www.entomol.org/journal/index.php?journal=JERS&page=article&op=viewArticle&path\[\]&pathID=341](http://www.entomol.org/journal/index.php?journal=JERS&page=article&op=viewArticle&path[]&pathID=341) [Accessed August 12, 2014].
- Bilski, M., 2012. Engineered genetic sterility of pest insects. University of Oxford.
- Bire, S. et al., 2013. Transposase concentration controls transposition activity: myth or reality? *Gene*, 530(2), pp.165–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23994686> [Accessed November 28, 2013].
- Black, M.J. et al., 2011. Life Cycle Assessment and sustainability methodologies for assessing industrial crops, processes and end products. *Industrial Crops and Products*, 34(2), pp.1332–1339. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0926669010003225> [Accessed October 30, 2011].
- Black, W.C., Alphey, L. & James, A. a, 2011. Why RIDL is not SIT. *Trends in parasitology*, 27(8), pp.362–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21659002> [Accessed August 21, 2011].
- Bloem, K.A., Bloem, S. & Carpenter, J.E., 2005. Impact of moth suppression/eradication programmes using the sterile insect technique or inherited sterility. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 677–700.

- Bloem, S. et al., 2004. Effect of Rearing Strategy and Gamma Radiation on Field Competitiveness of Mass-Reared Codling Moths (Lepidoptera : Tortricidae) Effect of Rearing Strategy and Gamma Radiation on Field Competitiveness of Mass-Reared Codling Moths (Lepidoptera : Tortric. *Journal of Economic Entomology*, 97, pp.1891–1898.
- Boake, C.R.B., Shelly, T.E. & Kaneshiro, K.Y., 1996. Sexual Selection in Relation to Pest-Management Strategies. *Annual review of entomology*, 41, pp.211–229.
- Bowman, A.W. & Azzalini, A., 1997. Applied smoothing techniques for data analysis: the kernel approach with S-plus illustrations, Oxford: Oxford University Press, Oxford, OX2 6DP.
- Brandt, K. et al., 2013. Methods for Comparing Data across Differently Designed Agronomic Studies: Examples of different Meta-analysis Methods Used to Compare Relative Composition of Plant Foods Grown Using Organic or Conventional Production Methods. *Journal of agricultural and food chemistry*, 61, pp.7173–7180.
- Le Bras, S. & Van Doren, M., 2006. Development of the male germline stem cell niche in *Drosophila*. *Developmental biology*, 294(1), pp.92–103. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16566915> [Accessed August 24, 2012].
- Brochta, D.A.O. & Atkinson, P.W., 1996. Mini-Review Transposable Elements and Gene Transformation in Non-Drosophilid Insects. *Insect biochemistry and molecular biology*, 26(96), pp.739–753.
- Brody, J.R. & Kern, S.E., 2004. History and principles of conductive media for standard DNA electrophoresis. *Analytical biochemistry*, 333(1), pp.1–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15351274> [Accessed July 12, 2012].
- Brown, S.J. et al., 2009. The red flour beetle, *Tribolium castaneum* (Coleoptera): a model for studies of development and pest biology. *Cold Spring Harbor protocols*, 2009(8), p.pdb.em0126. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20147228> [Accessed February 17, 2012].
- Buckman, K.A. et al., 2013. *Tribolium castaneum* (Coleoptera : Tenebrionidae) Associated with Rice Mills : Fumigation Efficacy and Population Rebound. *Journal of Economic Entomology*, 106(1), pp.499–512.
- Buja, A. et al., 2009. Statistical inference for exploratory data analysis and model diagnostics. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1906), pp.4361–83. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19805449> [Accessed March 6, 2013].
- Bulmer, M.G. & Parker, G. a, 2002. The evolution of anisogamy: a game-theoretic approach. *Proceedings. Biological sciences / The Royal Society*, 269(1507), pp.2381–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1691164&tool=pmcentrez&rendertype=abstract> [Accessed June 22, 2011].
- Burt, A., 2014. Heritable strategies for controlling insect vectors of disease Heritable strategies for controlling insect vectors of disease. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369.
- Calkins, C.O. & Parker, A.G., 2005. Sterile Insect Quality. In V. a. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 269–296.
- Carrasco Fernández, J.J. et al., 2012. MATLAB-based educational software for exploratory data analysis (EDA toolkit). *Computer Applications in Engineering Education*, 20(2), pp.313–320. Available at: <http://doi.wiley.com/10.1002/cae.20397> [Accessed April 24, 2013].

- Carrière, Y. et al., 2009. Cadherin gene expression and effects of Bt resistance on sperm transfer in pink bollworm. *Journal of insect physiology*, 55(11), pp.1058–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19666026> [Accessed July 19, 2012].
- Carson, R., 1962. *Silent Spring*, Houghton Mifflin, Boston, Massachusetts.
- Carvalho, D.O. et al., 2015. Suppression of a Field Population of Aedes aegypti in Brazil by Sustained Release of Transgenic Male Mosquitoes. *PLOS Neglected Tropical Diseases*, 9(7), p.e0003864. Available at: <http://dx.plos.org/10.1371/journal.pntd.0003864>.
- Carvalho, G.B., Ja, W.W. & Benzer, S., 2009. Non-lethal PCR genotyping of single Drosophila. *BioTechniques*, 46(4), pp.312–314.
- Castellanos, M.C. & Verdú, M., 2012. Meta-analysis of meta-analyses in plant evolutionary ecology. *Evolutionary Ecology*, 26(5), pp.1187–1196.
- Catteruccia, F., Benton, J.P. & Crisanti, A., 2005. An Anopheles transgenic sexing strain for vector control. *Nature biotechnology*, 23(11), pp.1414–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16244659> [Accessed October 7, 2011].
- Catteruccia, F., Godfray, H.C.J. & Crisanti, A., 2003. Impact of genetic manipulation on the fitness of Anopheles stephensi mosquitoes. *Science (New York, N.Y.)*, 299(5610), pp.1225–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12595691> [Accessed October 17, 2012].
- Chakraborty, S. et al., 2014. Vector modifications to eliminate transposase expression following piggyBac-mediated transgenesis. *Scientific reports*, 4, p.7403. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC41183/> [Accessed January 7, 2015].
- Chandran, D., Bergmann, F.T. & Sauro, H.M., 2009. TinkerCell: modular CAD tool for synthetic biology. *Journal of biological engineering*, 3, p.19. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC276589/> [Accessed February 28, 2013].
- Chapman, J.W. et al., 2002. High-altitude migration of the diamondback moth Plutella xylostella to the U.K.: a study using radar, aerial netting, and ground trapping. *Ecological Entomology*, 27(6), pp.641–650. Available at: <http://doi.wiley.com/10.1046/j.1365-2311.2002.00472.x>.
- Chapman, R.F., 2012. *The insects: structure and function* Fifth., Cambridge University Press.
- Charif, D. & Lobry, J.R., 2007. Structural approaches to sequence evolution: Molecules, networks, populations U. Bastolla et al., eds., Springer Verlag.
- Chen, C.Y. & Shyu, a B., 1995. AU-rich elements: characterization and importance in mRNA degradation. *Trends in biochemical sciences*, 20(11), pp.465–470.
- Clark, D.P. & Pazdernik, N.J., 2012. *Molecular Biology: Understanding the genetic revolution* Second., Oxford: Academic Press, Elsevier.
- Clarke, G. & Lunt, I., 2014. The concept of “originality” in the Ph.D.: how is it interpreted by examiners? *Assessment & Evaluation in Higher Education*, 39(7), pp.803–820. Available at: <http://www.tandfonline.com/doi/abs/10.1080/02602938.2013.870970>.
- Clopper, C. & Pearson, E., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26, pp.404–413.

- Concha, C. et al., 2011. Efficient germ-line transformation of the economically important pest species *Lucilia cuprina* and *Lucilia sericata* (Diptera, Calliphoridae). *Insect biochemistry and molecular biology*, 41(1), pp.70–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20869440> [Accessed May 2, 2013].
- Condon, K.C. et al., 2007. Germ-line transformation of the Mexican fruit fly. *Insect molecular biology*, 16(5), pp.573–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17894556>.
- Cooper, H. & Patall, E. a, 2009. The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological methods*, 14(2), pp.165–76. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19485627> [Accessed November 29, 2013].
- Cox, J.S.H. & Vreysen, M.J.B., 2005. Use of Geographic Information Systems and Spatial Analysis in Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 453–477.
- Cramer, J.S., 2003. The origins and development of the logit model. In *Logit Models from Economics and Other Fields*. Cambridge University Press, pp. 1–19.
- Crawley, M., 2005. *Statistics an introduction using R* Second., Chichester, West Sussex: Wiley & Sons.
- Cumming, G., Fidler, F. & Vaux, D.L., 2007. Error bars in experimental biology. *The Journal of cell biology*, 177(1), pp.7–11. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1837510/> [Accessed March 8, 2012].
- Cumming, G. & Finch, S., 2005. Inference by eye: confidence intervals and how to read pictures of data. *The American psychologist*, 60(2), pp.170–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15740449> [Accessed March 3, 2013].
- Dale, J. & Schantz, M., 2006. Analysis of Sequence Data. In *From Genes to Genomes*. John Wiley & Sons Ltd., Chichester, England, p. 196.
- Dallai, R., 2014. Overview on spermatogenesis and sperm structure of Hexapoda. *Arthropod Structure and Development*, 43(4), pp.257–290. Available at: <http://dx.doi.org/10.1016/j.asd.2014.04.002>.
- Daubenmire, R., 1936. The use of the terms syncytium and coenocyte in biology. *Science*, 84, p.533.
- Davies, J.P. & Grossman, a R., 1994. Sequences controlling transcription of the *Chlamydomonas reinhardtii* beta 2-tubulin gene after deflagellation and during the cell cycle. *Molecular and cellular biology*, 14(8), pp.5165–5174.
- Davis, M. & Miller, S., 1988. Expression of a testis β -tubulin gene in *Heliothis virescens* during spermatogenesis. *Insect biochemistry*, 18(4), pp.389–394. Available at: <http://www.sciencedirect.com/science/article/pii/0020179088900546> [Accessed October 24, 2013].
- Delattre, M., Tatout, C. & Coen, D., 2000. P-element transposition in *Drosophila melanogaster*: influence of size and arrangement in pairs. *Molecular and general genetics*, 263(3), pp.445–454.
- Denell, R., 2008. Establishment of *Tribolium* as a Genetic Model System and Its Early Contributions to Evo-Devo. *Genetics*, 180, pp.1779–1786. Available at:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448785/>&tool=pmcentrez&rendertype=abstract.

Dennett, D.C., 1995. *Darwin's Dangerous Idea*, Simon & Schuster.

Dorai-Raj, S., 2014. Package "binom": Binomial Confidence Intervals For Several Parameterizations. Available at: <http://cran.r-project.org/web/packages/binom/index.html>.

Dorogova, N. V., Nerusheva, O.O. & Omelyanchuk, L. V., 2009. Structural organization and dynamics of the endoplasmic reticulum during spermatogenesis of *Drosophila melanogaster*: Studies using PDI-GFP chimera protein. *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology*, 3(1), pp.55–61.

Downey, A., 2011. *Think Stats*, O'Reilly Media Inc.

Dransfield, R. & Brightwell, R., 2012. How to get on top of statistics: Design and Analysis for Biologists with R, InfluentialPoints, UK.

Dubben, H. & Beck-Bornholdt, H., 2005. Systematic review of publication bias in studies on publication bias. *BMJ (Clinical research ed.)*, 331(7514), pp.433–434. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1188108/>&tool=pmcentrez&rendertype=abstract [Accessed May 29, 2013].

Dubé, L., Pingali, P. & Webb, P., 2012. Paths of convergence for agriculture, health, and wealth. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31), pp.12294–301. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3412024/>&tool=pmcentrez&rendertype=abstract [Accessed October 4, 2012].

Durel, L., Estrada-peña, A. & Franc, M., 2015. Integrated fly management in European ruminant operations from the perspective of directive 2009/128/EC on sustainable use of pesticides. *Parasitology Research*, 114, pp.379–389.

Dutcher, S.K., 2001. Motile organelles: The importance of specific tubulin isoforms. *Current Biology*, 11(11), pp.419–422.

Dwan, K. et al., 2013. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One*, 8(7), p.e66844. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3702538/>&tool=pmcentrez&rendertype=abstract [Accessed November 4, 2013].

Dyck, V. a. et al., 2005. Management of Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 525–545.

Dyck, V.A., Hendrichs, J. & Robinson, A.S., 2005. *Sterile Insect Technique. Principles and Practice in Area-Wide Integrated Pest Management*. V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds., Springer.

Eddy, S.R., 1996. Hidden Markov models. *Current opinion in structural biology*, 6(3), pp.361–365.

Edvardsson, M. & Arnqvist, G., 2005. The effects of copulatory courtship on differential allocation in the red flour beetle *Tribolium castaneum*. *Journal of Insect Behavior*, 18(3), pp.313–322.

- El-Wakeil, N.E., 2010. Insect economic levels in relation to crop production. *Archives Of Phytopathology And Plant Protection*, 43(17), pp.1710–1744. Available at: <http://www.tandfonline.com/doi/abs/10.1080/03235400902753584> [Accessed January 10, 2012].
- Fackenthal, J., Turner, R. & Raff, E., 1993. Tissue-specific microtubule functions in Drosophila spermatogenesis requires the beta-2 tubulin isotype-specific carboxy terminus. *Developmental biology*, 158, pp.213–227.
- Fedina, T.Y. & Lewis, S.M., 2008. An integrative view of sexual selection in Tribolium flour beetles. *Biological reviews of the Cambridge Philosophical Society*, 83(2), pp.151–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18429767> [Accessed August 18, 2011].
- Ferguson, H.J. et al., 2011. Genetic transformation of the codling moth, Cydia pomonella L., with piggyBac EGFP. *Transgenic research*, 20(1), pp.201–14. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20386982> [Accessed May 2, 2013].
- Finch, S. & Cumming, G., 2009. Putting research in context: understanding confidence intervals from one or more studies. *Journal of pediatric psychology*, 34(9), pp.903–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19095715> [Accessed April 9, 2013].
- Fire, A. et al., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391(February), pp.806–811.
- Fleming, C. & O'Carroll, J., 2011. Revolution, rupture, rhetoric. *Philosophy & Social Criticism*, 38(1), pp.39–57. Available at: <http://psc.sagepub.com/cgi/doi/10.1177/0191453711421603> [Accessed November 1, 2012].
- Folcher, M. et al., 2013. Synthetic mammalian trigger-controlled bipartite transcription factors. *Nucleic Acids Research*, 41(13).
- França, L.T.C., Carrilho, E. & Kist, T.B.L., 2002. A review of DNA sequencing techniques. *Quarterly reviews of biophysics*, 35(2), pp.169–200. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12197303>.
- Franz, A.W.E. et al., 2011. Comparison of transgene expression in *Aedes aegypti* generated by mariner Mos1 transposition and ΦC31 site-directed recombination. *Insect molecular biology*, 20(5), pp.587–98. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21699593> [Accessed February 21, 2012].
- Fraser, M.J. et al., 1983. Acquisition of host cell DNA sequences by Baculoviruses : relationship between host mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. *Journal of Virology*, 47, pp.187–200.
- Fraser, M.J., 2012. Insect transgenesis: current applications and future prospects. *Annual review of entomology*, 57, pp.267–89. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22149266> [Accessed August 31, 2012].
- Friedländer, M., 1997. Control of the eupyrene-apyrene sperm dimorphism in Lepidoptera. *Journal of insect physiology*, 43(12), pp.1085–1092. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12770480>.
- Fu, G. et al., 2007. Female-specific insect lethality engineered using alternative splicing. *Nature biotechnology*, 25(3), pp.353–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17322873> [Accessed September 13, 2011].

- Gallia, G.L. & Khalili, K., 1998. Evaluation of an autoregulatory tetracycline regulated system. *Oncogene*, 16(14), pp.1879–1884.
- Garcia-Perez, M., 2005. On the confidence interval for the binomial parameter. *Quality and Quantity*, 39(4), pp.467–481.
- Gerland, P. et al., 2014. World population stabilization unlikely this century. *Science*, 346(6206), pp.234–237.
- Gitzinger, M. et al., 2012. The food additive vanillic acid controls transgene expression in mammalian cells and mice. *Nucleic Acids Research*, 40(5).
- Godfray, H.C.J. et al., 2010. Food security: the challenge of feeding 9 billion people. *Science (New York, N.Y.)*, 327(5967), pp.812–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20110467> [Accessed July 26, 2011].
- Gong, P. et al., 2005. A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. *Nature biotechnology*, 23(4), pp.453–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15750586> [Accessed October 19, 2011].
- Goshima, G. et al., 2007. Genes required for mitotic spindle assembly in Drosophila S2 cells. *Science (New York, N.Y.)*, 316(APRIL), pp.417–421.
- Gossen, M. & Bujard, H., 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), pp.5547–51. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC40130/> [Accessed November 7, 2013].
- Grace, N.D., 1966. The present status of shunts for portal hypertension in cirrhosis. *Journal of Gastroenterology*, 50(5), pp.253–272.
- Grazer, V.M. & Martin, O.Y., 2011. Elevated temperature changes female costs and benefits of reproduction. *Evolutionary Ecology*, 26(3), pp.625–637. Available at: <http://www.springerlink.com/index/10.1007/s10682-011-9508-4.html> [Accessed October 17, 2012].
- Green, R.E. et al., 2005. Farming and the fate of wild nature. *Science (New York, N.Y.)*, 307(5709), pp.550–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15618485> [Accessed November 7, 2013].
- Grolemund, G. & Wickham, H., 2014. A Cognitive Interpretation of Data Analysis. *International Statistical Review*, pp.1–25. Available at: <http://onlinelibrary.wiley.com/doi/10.1111/insr.12028/full> [Accessed October 1, 2014].
- Gross, L. a et al., 2000. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proceedings of the National Academy of Sciences of the United States of America*, 97, pp.11990–11995.
- Grossman, G.L. et al., 2001. Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. *Insect molecular biology*, 10(6), pp.597–604. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11903629>.
- Haddad, N.M. et al., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advancements*, 1(e1500052), pp.1–9.

- Hand, D.J., 1998. Data Mining : Statistics and More ? *The American Statistician*, 52(2), pp.112–118.
- Handel, K. et al., 2000. Tribolium embryogenesis: a SEM study of cell shapes and movements from blastoderm to serosal closure. *Development genes and evolution*, 210(4), pp.167–79. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11180819>.
- Handler, a M. et al., 1998. The lepidopteran transposon vector, piggyBac, mediates germ-line transformation in the Mediterranean fruit fly. *Proceedings of the National Academy of Sciences of the United States of America*, 95(13), pp.7520–7525.
- Handler, a M. & Harrell, R. a, 2001. Transformation of the Caribbean fruit fly, *Anastrepha suspensa*, with a piggyBac vector marked with polyubiquitin-regulated GFP. *Insect biochemistry and molecular biology*, 31(2), pp.199–205. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11164342>.
- Handler, A.M., 2000. An introduction to the history and methodology of insect gene transfer. In A. M. Handler & A. a James, eds. *Insect transgenesis methods and applications*. CRC Press LLC, Boca Raton, Florida.
- Handler, A.M. & Beeman, R.W., 2003. United States Department of Agriculture?Agricultural Research Service: advances in the molecular genetic analysis of insects and their application to pest management. *Pest Management Science*, 59(6-7), pp.728–735. Available at: <http://doi.wiley.com/10.1002/ps.719> [Accessed August 28, 2012].
- Handler, A.M. & Harrell, R.A., 1999. Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. *Insect molecular biology*, 8(4), pp.449–57. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10634970>.
- Hanley, J.A. & Lippman-Hand, A., 1983. If nothing goes wrong, is everything all right? *JAMA: the journal of the American ...*, 249(13), pp.1743–1745. Available at: <http://jama.ama-assn.org/content/249/13/1743.short> [Accessed August 3, 2012].
- Happ, G., 1992. Maturation of the male reproductive system and its endocrine regulation. *Annual review of entomology*, 37(9), pp.303–320. Available at: <http://www.annualreviews.org/doi/pdf/10.1146/annurev.en.37.010192.001511> [Accessed July 22, 2014].
- Hara, K. et al., 2009. A novel fusion protein that functions as an enhanced green fluorescent protein reporter and a tetracycline-controlled transcriptional activator. *Development Genes and Evolution*, 219(2), pp.103–110. Available at: <http://www.springerlink.com/index/10.1007/s00427-009-0274-z> [Accessed May 8, 2012].
- Hargreaves, C. & Francis, B., 2013. The long term recidivism risk of young sexual offenders in England and Wales—enduring risk or redemption? *Journal of Criminal Justice*, pp.1–30. Available at: <http://www.sciencedirect.com/science/article/pii/S0047235213000561> [Accessed July 29, 2013].
- Harris, A.F. et al., 2012. Successful suppression of a field mosquito population by sustained release of engineered male mosquitoes. *Nature Biotechnology*, 30(9), pp.828–830.
- Harvey-Samuel, T., 2014. Genetic Control of the Diamondback moth (*Plutella xylostella* L.). The University of Oxford.
- Harvey-Samuel, T. et al., 2014. Population-level effects of fitness costs associated with repressible female-lethal transgene insertions in two pest insects. *Evolutionary Applications*, 7(5), pp.597–606.

- Hawk, J.A., Frey, J. & Bell, J., 1972. Photic effects on the fecundity of *Tribolium castaneum*, a preliminary report. *Tribolium Inf. Bull.*, 15, pp.77–79.
- Headey, D. & Fan, S., 2008. Anatomy of a crisis: the causes and consequences of surging food prices. *Agricultural Economics*, 39, pp.375–391. Available at: <http://doi.wiley.com/10.1111/j.1574-0862.2008.00345.x> [Accessed July 21, 2011].
- Hediger, M. et al., 2001. Genetic transformation of the housefly *Musca domestica* with the lepidopteran derived transposon piggyBac. *Insect molecular biology*, 10(2), pp.113–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11422506>.
- Heinrich, J.C. et al., 2002. Germ-line transformation of the Australian sheep blowfly *Lucilia cuprina*. *Insect molecular biology*, 11(1), pp.1–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11841497>.
- Helinski, M.E.H., Parker, A.G. & Knols, B.G.J., 2009. Radiation biology of mosquitoes. *Malaria journal*, 8 Suppl 2, p.S6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2777328&tool=pmcentrez&rendertype=abstract> [Accessed November 18, 2011].
- Hendrichs, J. et al., 2005. Strategic Options in using Sterile Insects for Area-Wide Integrated Pest Management. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 563–600.
- Hillis, W.D. & Daniel Hillis, W., 1993. Why physicists like models and why biologists should. *Current Biology*, 3(2), pp.79–81. Available at: <http://linkinghub.elsevier.com/retrieve/pii/096098229390159L>.
- Hokanson, K.E. et al., 2013. Not all GMOs are crop plants: non-plant GMO applications in agriculture. *Transgenic Research*. Available at: <http://link.springer.com/10.1007/s11248-013-9769-5> [Accessed November 19, 2013].
- Hope, A., 1968. A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical society B*, 30, pp.582–598.
- Horn, C. et al., 2002. Fluorescent transformation markers for insect transgenesis. *Insect biochemistry and molecular biology*, 32(10), pp.1221–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12225913>.
- Howe, R.W., 1956. The effect of temperature and humidity on the rate of development and mortality of *Tribolium Castaneum* (Herbst). *Annals of Applied Biology*, 44, pp.356–368.
- Howe, R.W., 1956. The effects of temperature and humidity on the oviposition rate of *Tribolium castaneum*. *Bulletin of Entomological Research*, 53, pp.301–310.
- Howell, D.C., 2010. CHI-SQUARE TEST - ANALYSIS OF CONTINGENCY TABLES. In M. Lovric, ed. *International Encyclopedia of Statistical Science*. Berlin, pp. 1–4. Available at: <http://www.uvm.edu/~dhowell/methods7/Supplements/ChiSquareTests.pdf>.
- Hoyle, H.D. et al., 1995. Regulation of beta-tubulin function and expression in *Drosophila* spermatogenesis. *Developmental genetics*, 16(2), pp.148–170.
- Huf, W. et al., 2011. Meta-analysis: fact or fiction? How to interpret meta-analyses. *The world journal of biological psychiatry : the official journal of the World Federation of Societies of Biological Psychiatry*, 12(3), pp.188–200. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21375365> [Accessed November 21, 2013].

- Hurlbert, S.H., 2009. The ancient black art and transdisciplinary extent of pseudoreplication. *Journal of comparative psychology (Washington, D.C. : 1983)*, 123(4), pp.434–43. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19929111> [Accessed October 4, 2012].
- Huvenne, H. & Smagghe, G., 2010. Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: a review. *Journal of insect physiology*, 56(3), pp.227–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19837076> [Accessed August 11, 2014].
- Ihaka, R. & Gentleman, R., 1996. R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, 5(3), pp.299–314.
- Ishii, G. & Hayakawa, R., 1960. On the compound binomial distribution. *Annals of the institute of statistical mathematics*, 12(1), pp.69–80.
- Jackson, R.J., Hellen, C.U.T. & Pestova, T. V., 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews. Molecular cell biology*, 11(2), pp.113–127. Available at: <http://dx.doi.org/10.1038/nrm2838>.
- Jaeger, T.F., 2008. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language*, 59(4), pp.434–446. Available at: <http://www.sciencedirect.com/science/article/pii/S0749596X07001337> [Accessed May 7, 2013].
- Jaynes, E.T., 1976. Confidence Intervals vs. Bayesian Intervals. In W. L. Harper & C. A. Hooker, eds. *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Dordrecht: D. Reidel, pp. 175–257.
- Jin, L., 2011. Development of a sperm marker and a sperm-lethal transgenic strain in the Mediterranean fruit fly , *Ceratitis capitata*. The University of Oxford.
- Jones, A.M., Ehrhardt, D.W. & Frommer, W.B., 2012. A never ending race for new and improved fluorescent proteins. *BMC biology*, 10, p.39. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3342923/> [tool=pmcentrez&rendertype=abstract].
- Jones, O., Maillardet, R. & Robinson, A., 2009. *Introduction to Scientific Programming and Simulation using R*, Boca Raton, Florida: Chapman & Hall.
- Justus, K.A. & Mitchell, B.K., 1999. Reproductive morphology, copulation, and inter-populational variation in the diamondback moth, *Plutella xylostella* (L.) (Lepidoptera: Pluellidae). *International Journal of Insect Morphology and Embryology*, 28, pp.233–246.
- Kaczorowski, T., Skowron, P. & Podhajska, a J., 1989. Purification and characterization of the FokI restriction endonuclease. *Gene*, 80(2), pp.209–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2583511>.
- Kanippayoor, R.L., Alpern, J.H.M. & Moehring, A.J., 2013. Protamines and spermatogenesis in *Drosophila* and *Homo sapiens* : A comparative analysis. *Spermatogenesis*, 3(2), p.e24376. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3710222/> [tool=pmcentrez&rendertype=abstract].
- Kapetanaki, M.G. et al., 2002. High frequencies of Minos transposon mobilization are obtained in insects by using in vitro synthesized mRNA as a source of transposase. *Nucleic acids research*, 30(15), pp.3333–40. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC137079/> [tool=pmcentrez&rendertype=abstract].

- Kasinsky, H., Eirin-lopez, J. & Ausio, J., 2011. Protamines: Structural complexity, evolution and chromatin patterning. *Protein and peptide letters*, 18(8), pp.755–771.
- Kawasaki, H. et al., 2003. Analysis of α - and β -tubulin genes of *Bombyx mori* using an EST. *Insect biochemistry and molecular biology*, 33, pp.131–137.
- Kemphues, K.J. et al., 1982. The testis-specific beta-tubulin subunit in *Drosophila melanogaster* has multiple functions in spermatogenesis. *Cell*, 31(3 Pt 2), pp.655–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6819086>.
- Khoshdel, a, Attia, J. & Carney, S.L., 2006. Basic concepts in meta-analysis: A primer for clinicians. *International journal of clinical practice*, 60(10), pp.1287–94. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16981972> [Accessed September 17, 2013].
- Kim, A. & Pyykko, I., 2011. Size matters: versatile use of PiggyBac transposons as a genetic manipulation tool. *Molecular and cellular biochemistry*, 354(1-2), pp.301–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21516337> [Accessed July 7, 2011].
- Kim, H. et al., 2012. Mutagenesis by imprecise excision of the piggyBac transposon in *Drosophila melanogaster*. *Biochemical and Biophysical Research Communications*, 417, pp.335–339.
- Kimanani, E.K. & Odhiambo, J., 1993. Model for the Sterile Male Technique. *Biometrics*, 49(2), pp.577–585.
- Kitano, H., 2002. Systems biology: a brief overview. *Science (New York, N.Y.)*, 295(5560), pp.1662–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11872829> [Accessed May 23, 2013].
- Klassen, W. & Curtis, C.F., 2005. History of the Sterile Insect Technique. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 3–36.
- Knippling, E.F. & Serv, A.R., 1955. Possibilities of Insect Control or Eradication Through the Use of Sexually Sterile Males. *Journal of Economic Entomology*, (August), pp.459–462.
- Kotwica, J. et al., 2011. Diurnal rhythm in expression and release of yolk protein in the testis of *Spodoptera littoralis* (Lepidoptera: Noctuidae). *Insect biochemistry and molecular biology*, 41(4), pp.264–72. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21216288> [Accessed August 27, 2014].
- Krafsur, E.S., 1998. Sterile insect technique for suppressing and eradicating insect populations: 55 years and counting. *Journal of agricultural entomology*, 15(4), pp.303–317.
- Kreutz, C. & Timmer, J., 2009. Systems biology: experimental design. *The FEBS journal*, 276(4), pp.923–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19215298> [Accessed March 11, 2013].
- Kuersten, S. & Goodwin, E.B., 2003. The power of the 3' UTR: translational control and development. *Nature reviews. Genetics*, 4(8), pp.626–637.
- Kuwayama, H. et al., 2006. Germ-line transformation and RNAi of the ladybird beetle, *Harmonia axyridis*. *Insect molecular biology*, 15(4), pp.507–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16907837>.
- Labbé, G.M.C. et al., 2012. Female-specific flightless (fsRIDL) phenotype for control of *Aedes albopictus*. *PLoS neglected tropical diseases*, 6(7), p.e1724. Available at:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3393675/>&tool=pmcentrez&rendertype=abstract [Accessed November 28, 2012].

Labbé, G.M.C., Nimmo, D.D. & Alphey, L., 2010. piggybac- and PhiC31-mediated genetic transformation of the Asian tiger mosquito, *Aedes albopictus* (Skuse). *PLoS neglected tropical diseases*, 4(8), p.e788. Available at:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2923142/>&tool=pmcentrez&rendertype=abstract [Accessed August 16, 2011].

Ladyman, J., 2002. *Understanding Philosophy of Science*, Milton Park, Abingdon, Oxfordshire: Routledge.

Lagisz, M., Wilde, K.E. & Wolff, K., 2010. The development of PCR-based markers for molecular sex identification in the model insect species *Tribolium castaneum*. *Entomologia Experimentalis et Applicata*, 134(1), pp.50–59. Available at: <http://doi.wiley.com/10.1111/j.1570-7458.2009.00935.x> [Accessed August 29, 2013].

Leach, a W. & Mumford, J.D., 2008. Pesticide Environmental Accounting: a method for assessing the external costs of individual pesticide applications. *Environmental pollution (Barking, Essex : 1987)*, 151(1), pp.139–47. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17604888> [Accessed October 31, 2011].

Van Lenteren, J.C., 2000. A greenhouse without pesticides: Fact or fantasy? *Crop Protection*, 19(6), pp.375–384.

Lerer-Goldshtein, T. et al., 2010. TMF/ARA160: A key regulator of sperm development. *Developmental biology*, 348(1), pp.12–21. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/20691678> [Accessed July 19, 2012].

Levine, M. & Tjian, R., 2003. Transcription regulation and animal diversity. *Nature*, 424(6945), pp.147–51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12853946>.

Lewis, S. & Clarke, M., 2001. Forest plots: trying to see the wood and the trees. *BMJ (Clinical research ed.)*, 322(7300), pp.1479–80. Available at:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1120528/>&tool=pmcentrez&rendertype=abstract.

Li, J. & Biggin, M., 2015. Statistics requantitates the central dogma. *Science*, 347(6226), pp.1066–1067.

Li, K. et al., 1998. Drosophila centrosomin protein is required for male meiosis and assembly of the flagellar axoneme. *The Journal of cell biology*, 141(2), pp.455–67. Available at:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2148447/>&tool=pmcentrez&rendertype=abstract.

Li, X., Heinrich, J.C. & Scott, M.J., 2001. piggyBac-mediated transposition in *Drosophila melanogaster*: an evaluation of the use of constitutive promoters to control transposase gene expression. *Insect molecular biology*, 5(5), pp.447–455. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/11881809>.

Limpert, E. & Stahel, W. a, 2011. Problems with using the normal distribution--and ways to improve quality and efficiency of data analysis. *PloS one*, 6(7), p.e21403. Available at:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136454/>&tool=pmcentrez&rendertype=abstract [Accessed July 19, 2012].

Lin, C.-Y. & Yang, M.-C., 2008. Improved p -Value Tests for Comparing Two Independent Binomial Proportions. *Communications in Statistics - Simulation and Computation*, 38(1), pp.78–

91. Available at: <http://www.tandfonline.com/doi/abs/10.1080/03610910802417812> [Accessed April 17, 2013].
- Lin, H. et al., 2007. Cellular toxicity induced by SRF-mediated transcriptional squelching. *Toxicological sciences : an official journal of the Society of Toxicology*, 96(1), pp.83–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17116645> [Accessed August 22, 2012].
- Little, R. & Rubin, D.B., 2002. *Statistical Analysis with Missing Data*, John Wiley & Sons Ltd., Hoboken, New Jersey.
- Liu, D. et al., 2012. Genetic transformation mediated by piggyBac in the Asian corn borer, *Ostrinia furnacalis* (Lepidoptera: Crambidae). *Archives of insect biochemistry and physiology*, 80(3), pp.140–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22696097> [Accessed August 23, 2012].
- Livermore, D.M., 1995. beta-Lactamases in laboratory and clinical resistance. *Clinical microbiology reviews*, 8(4), pp.557–84. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC172876/> [tool=pmcentrez&rendertype=abstract].
- Lobo, N.F. et al., 2002. Germ line transformation of the yellow fever mosquito, *Aedes aegypti*, mediated by transpositional insertion of a piggyBac vector. *Insect molecular biology*, 11(2), pp.133–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11966878>.
- Lorenzen, M.D. et al., 2007. piggyBac-based insertional mutagenesis in *Tribolium castaneum* using donor/helper hybrids. *Insect molecular biology*, 16(3), pp.265–75. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17316329> [Accessed July 23, 2011].
- Lorenzen, M.D. et al., 2003. piggyBac-mediated germline transformation in the beetle *Tribolium castaneum*. *Insect molecular biology*, 12(5), pp.433–40. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12974948>.
- Lorenzen, M.D. et al., 2002. Transgene expression from the *Tribolium castaneum* Polyubiquitin promoter. *Insect molecular biology*, 11(5), pp.399–407. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18625397>.
- Lu, J., Maddison, L. a & Chen, W., 2011. PhiC31 integrase induces efficient site-specific excision in zebrafish. *Transgenic research*, 20(1), pp.183–9. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3019273/> [tool=pmcentrez&rendertype=abstract] [Accessed March 4, 2012].
- Ludbrook, J. & Lew, M.J., 2009. Estimating the risk of rare complications: is the “rule of three” good enough? *ANZ journal of surgery*, 79(7-8), pp.565–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19694672> [Accessed August 20, 2012].
- Lukacs, G. et al., 2000. Size dependent DNA mobility in cytoplasm and nucleus. *J Biol Chem*, 275, pp.1625–1629.
- Lutz, W. et al., 2014. Population growth : Peak probability and Limits of food supply. *Science*, 346(6209), p.560.
- Lydersen, S., Fagerland, M.W. & Laake, P., 2009. Recommended tests for association in 2×2 tables. *Statistics in medicine*, 28, pp.1159–1175.
- Malthus, T., 1798. An essay on the principle of population. In *Oxford World's Classics*. Oxford: Oxford University Press, Oxford, OX2 6DP.

- Mandrioli, M. & Wimmer, E. a, 2003. Stable transformation of a *Mamestra brassicae* (lepidoptera) cell line with the lepidopteran-derived transposon piggyBac. *Insect biochemistry and molecular biology*, 33(1), pp.1–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12459194>.
- Mao, Y.-B. et al., 2007. Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nature biotechnology*, 25(11), pp.1307–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17982444> [Accessed November 2, 2012].
- Marec, F., Kollarova, I. & Pavelka, J., 1999. Radiation-Induced Sterility Combined with a Genetic Sexing System in *Ephestia kuehniella* (Lepidoptera: Pyralidae). *Annals of the Entomological Society of America*, 92(2), pp.251–259.
- Marec, F., Neven, L. & Robinson, A., 2005. Development of genetic sexing strains in Lepidoptera: from traditional to transgenic approaches. *Journal of Economic Entomology*, 98(2), pp.248–259. Available at: <http://www.bioone.org/doi/abs/10.1603/0022-0493-98.2.248> [Accessed November 1, 2013].
- Marrelli, M.T. et al., 2006. Mosquito transgenesis: what is the fitness cost? *Trends in parasitology*, 22(5), pp.197–202. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16564223> [Accessed July 15, 2011].
- Marrelli, M.T. et al., 2007. Transgenic malaria-resistant mosquitoes have a fitness advantage when feeding on Plasmodium-infected blood. *Proceedings of the National Academy of Sciences of the United States of America*, 104(13), pp.5580–5583.
- Martins, S. et al., 2012. Germline transformation of the diamondback moth, *Plutella xylostella* L., using the piggyBac transposable element. *Insect Molecular Biology*, 44(0), p.no-no. Available at: <http://doi.wiley.com/10.1111/j.1365-2583.2012.01146.x> [Accessed May 24, 2012].
- Matthews, G.A., 2006. *Pesticides: Health, Safety and the Environment*, Blackwell Publishing, Oxford.
- Maxmen, A., 2013. Under attack. *Nature*, 501, pp.5–7.
- Mazumder, B., Seshadri, V. & Fox, P.L., 2003. Translational control by the 3'-UTR: The ends specify the means. *Trends in Biochemical Sciences*, 28(2), pp.91–98.
- McClain, C.R. et al., 2015. Sizing ocean giants: patterns of intraspecific size variation in marine megafauna. *PeerJ*, 2, p.e715. Available at: <https://peerj.com/articles/715> [Accessed January 13, 2015].
- McClintock, B., 1950. The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America*, 36, pp.344–355.
- McGraw, E. a & O'Neill, S.L., 2013. Beyond insecticides: new thinking on an ancient problem. *Nature reviews. Microbiology*, 11(3), pp.181–93. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23411863> [Accessed March 2, 2013].
- McShane, B.B. & Bockenholt, U., 2013. The Power of Weak Studies : Why the Synthesis of a Research Paper Matters. *Advances in consumer research*, 41, pp.505–506.
- Mehier-Humbert, S. & Guy, R.H., 2005. Physical methods for gene transfer: improving the kinetics of gene delivery into cells. *Advanced drug delivery reviews*, 57(5), pp.733–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15757758> [Accessed November 6, 2012].

- Michalova, E., Vojtesek, B. & Hrstka, R., 2013. Impaired pre-mRNA processing and altered architecture of 3' untranslated regions contribute to the development of human disorders. *International Journal of Molecular Sciences*, 14(8), pp.15681–15694.
- Mickel, C.E. & Standish, J., 1947. Susceptibility of processed soy flour and soy grits in storage to attack by *Tribolium castaneum* (Herbst). *J. Dep. Agric. Vict.*, 42, pp.365–373.
- Miller, J.C. et al., 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature biotechnology*, 25(7), pp.778–85. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17603475> [Accessed July 26, 2011].
- Miller, S.C. et al., 2012. Dissecting Systemic RNA Interference in the Red Flour Beetle *Tribolium castaneum*: Parameters Affecting the Efficiency of RNAi. *PLoS ONE*, 7(10).
- Misof, B. et al., 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), pp.763–768.
- Mita, K. & Nenoi, M., 1995. Expression of the *Bombyx mori* beta-tubulin-encoding gene in testis. *Gene*, 162, pp.329–330.
- Mitra, R., Fain-Thornton, J. & Craig, N.L., 2008. piggyBac can bypass DNA synthesis during cut and paste transposition. *The EMBO journal*, 27(7), pp.1097–109. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2323262&tool=pmcentrez&rendertype=abstract> [Accessed October 14, 2011].
- Mo, J. et al., 2003. Local Dispersal of the Diamondback Moth (*Plutella xylostella* (L.)) (Lepidoptera: Plutellidae). *Environmental Entomology*, 32(1), pp.71–79. Available at: <http://www.bioone.org/doi/abs/10.1603/0046-225X-32.1.71> [Accessed November 1, 2011].
- Mohammed, A. & Coates, C.J., 2004. Promoter and piggyBac activities within embryos of the potato tuber moth, *Phthorimaea operculella*, Zeller (Lepidoptera: Gelechiidae). *Gene*, 342(2), pp.293–301. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15527988> [Accessed March 20, 2013].
- Mon, H. et al., 2011. Double-strand breaks repair by gene conversion in silkworm holocentric chromosomes. *Molecular genetics and genomics : MGG*, pp.215–224. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21842267> [Accessed August 29, 2011].
- Morales, M.E. et al., 2007. piggyBac transposon mediated transgenesis of the human blood fluke, *Schistosoma mansoni*. *The FASEB journal official publication of the Federation of American Societies for Experimental Biology*, 21, pp.3479–3489.
- Morange, M., 2009. What history tells us XIX. The notion of the episome. *Journal of Biosciences*, 34(6), pp.845–848.
- Morey, R. et al., 2015. The Fallacy of Placing Confidence in Confidence Intervals. *Psychonomic Bulletin & Review*, in press.
- Morris, a C., Eggleston, P. & Crampton, J.M., 1989. Genetic transformation of the mosquito *Aedes aegypti* by micro-injection of DNA. *Medical and veterinary entomology*, 3(1), pp.1–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2519641>.
- Morris, C. a, Benson, E. & White-Cooper, H., 2009. Determination of gene expression patterns using in situ hybridization to *Drosophila* testes. *Nature protocols*, 4(12), pp.1807–19. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20010932> [Accessed April 13, 2012].

- Morris, E.J., 2011. A semi-quantitative approach to GMO risk-benefit analysis. *Transgenic research*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21197601> [Accessed August 23, 2011].
- Morrison, N.I. et al., 2012. Engineered Repressible Lethality for Controlling the Pink Bollworm , a Lepidopteran Pest of Cotton. *PloS one*, 7(12), pp.1–10.
- Morrison, N.I., *Genetic control of agricultural pests*. University of Oxford.
- Morrison, N.I. et al., 2010. Genetic Improvements to the Sterile Insect Technique for Agricultural Pests. *AsPac J. Mol. Biotechnol.*, 18(2), pp.275–295.
- Mumford, J.D., 2005. Application of Benefit/Cost Analysis to Insect Pest Control using the Sterile Insect Technique. In V. A. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 481–498.
- Murakami, a & Imai, H.T., 1974. Cytological evidence for holocentric chromosomes of the silkworms, *Bombyx mori* and *B. mandarina*, (Bombycidae, Lepidoptera). *Chromosoma*, 47(2), pp.167–78. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4141955>.
- Nakanishi, H. et al., 2011. Comparison of piggyBac transposition efficiency between linear and circular donor vectors in mammalian cells. *Journal of biotechnology*, 154(4), pp.205–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21640767> [Accessed April 9, 2013].
- Nei, M. & Saitou, N., 1987. The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular biology and evolution*, 4(4), pp.406–425.
- Newcombe, R.G., 1998a. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Statistics in medicine*, 17(8), pp.873–90. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9595617>.
- Newcombe, R.G., 1998b. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in medicine*, 17(8), pp.857–72. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9595616>.
- Nielsen, M.G. et al., 2001. Axoneme-specific beta-tubulin specialization: a conserved C-terminal motif specifies the central pair. *Current biology : CB*, 11(7), pp.529–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11413005>.
- Nielsen, M.G. et al., 2006. Functional constraint underlies 60 million year stasis of Dipteran testis-specific B2-tubulin. *Evolution and Development*, 8(1), pp.23–29.
- Nielsen, M.G., Gadagkar, S.R. & Gutzwiller, L., 2010. Tubulin evolution in insects: gene duplication and subfunctionalization provide specialized isoforms in a functionally constrained gene family. *BMC evolutionary biology*, 10, p.113. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2880298&tool=pmcentrez&rendertype=abstract>.
- Nimmo, D.D. et al., 2006. High efficiency site-specific genetic engineering of the mosquito genome. *Insect molecular biology*, 15(2), pp.129–36. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1602059&tool=pmcentrez&rendertype=abstract> [Accessed August 16, 2011].
- O'Brochta, D. a., 2003. Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology*, 206(21), pp.3823–3834. Available at: <http://jeb.biologists.org/cgi/doi/10.1242/jeb.00638> [Accessed June 27, 2011].

- O'Donnell, K.H., Chen, C.T. & Wensink, P.C., 1994. Insulating DNA directs ubiquitous transcription of the *Drosophila melanogaster* alpha 1-tubulin gene. *Molecular and cellular biology*, 14(9), pp.6398–6408.
- O'Donnell, L. & O'Bryan, M.K., 2014. Microtubules and spermatogenesis. *Seminars in Cell and Developmental Biology*, 30, pp.45–54. Available at: <http://dx.doi.org/10.1016/j.semcd.2014.01.003>.
- O'Hagan, A., 2009. Bayesian principles. In *Bayesian Methods in Health Economics: a short course, The Biomedical & Life Sciences Collection*. London. Available at: <http://ezproxy.ouls.ox.ac.uk:4492/bio>.
- O'Hagan, A. & Luce, B., 2003. *A primer on Bayesian statistics in health economics and outcomes research*, MEDTAP International, Inc. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:a+primer+on+Bayesian+statistics+in+health+economics+and+outcomes+research#0> [Accessed May 14, 2013].
- Oppenheim, S.J. et al., 2015. We can't all be supermodels: the value of comparative transcriptomics to the study of non-model insects. *Insect Molecular Biology*, 24(2014), p.n/a-n/a. Available at: <http://doi.wiley.com/10.1111/imb.12154>.
- Ostermeier, M., 2007. Beyond Cataloging the Library of Babel. *Chemistry and Biology*, 14(3), pp.237–238.
- Pai, A. & Bernasconi, G., 2008. Polyandry and Female Control : The Red Flour Beetle *Tribolium castaneum* as a Case Study. *Science*, 159(September 2006), pp.148–159.
- Papathanos, P. a et al., 2009. Sex separation strategies: past experience and new approaches. *Malaria journal*, 8 Suppl 2, p.S5. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2777327/> [Accessed July 28, 2011].
- Park, T., 1934. Observations on the General Biology of the Flour Beetle. *Quarterly Review of Biology*, 9(1), pp.36–54.
- Patterson, G., Day, R. & Piston, D., 2001. Fluorescent protein spectra. *Journal of Cell Science*, 114, pp.837–838. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2220689/>.
- Pavlopoulos, A. et al., 2004. Efficient transformation of the beetle *Tribolium castaneum* using the Minos transposable element: quantitative and qualitative analysis of genomic integration events. *Genetics*, 167(2), pp.737–746. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1470898/> [Accessed July 28, 2011].
- Pease, J.B. & Hahn, M.W., 2012. Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Molecular Biology and Evolution*, 29(6), pp.1645–1653.
- Perera, O.P., Harrell II, R. a & Handler, a M., 2002. Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a piggyBac/EGFP transposon vector is routine and highly efficient. *Insect molecular biology*, 11(4), pp.291–7. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1214469/>.
- Perez-Mendoza, J., Campbell, J.F. & Throne, J.E., 2011. Effects of Rearing Density, Age, Sex, and Food Deprivation on Flight Initiation of the Red Flour Beetle (Coleoptera: Tenebrionidae). *Journal of Economic Entomology*, 104(2), pp.443–451. Available at: <http://www.bioone.org/doi/abs/10.1603/EC10430> [Accessed October 18, 2012].

- Pérez-Staples, D., Shelly, T.E. & Yuval, B., 2013. Female mating failure and the failure of "mating" in sterile insect programs. *Entomologia Experimentalis et Applicata*, 146(1), pp.66–78. Available at: <http://doi.wiley.com/10.1111/j.1570-7458.2012.01312.x> [Accessed August 8, 2014].
- Peters, J. et al., 2006. Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *JAMA: the journal of the American ...*, 295(6), pp.676–680. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Publication+Bias+in+Meta-analysis#0> [Accessed December 2, 2013].
- Petitti, D., 2004. Commentary: hormone replacement therapy and coronary heart disease: four lessons. *International journal of epidemiology*, 33(3), pp.461–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15166191> [Accessed November 19, 2012].
- Pfeiffer, B.D. et al., 2010. Refinement of tools for targeted gene expression in *Drosophila*. *Genetics*, 186(2), pp.735–55. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2942869&tool=pmcentrez&rendertype=abstract> [Accessed March 6, 2012].
- Phuc, H.K. et al., 2007. Late-acting dominant lethal genetic systems and mosquito control. *BMC biology*, 5, p.11. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1865532&tool=pmcentrez&rendertype=abstract> [Accessed July 21, 2014].
- Piégu, B. et al., 2015. A survey of transposable element classification systems – A call for a fundamental update to meet the challenge of their diversity and complexity. *Molecular Phylogenetics and Evolution*, 86, pp.90–109. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1055790315000664>.
- Pigliucci, M. & Kaplan, J., 2000. The fall and rise of Dr Pangloss: Adaptationism and the Spandrels paper 20 years later. *Trends in Ecology and Evolution*, 15(2), pp.66–70.
- Pimbert, M., 1991. Designing integrated pest management for sustainable and productive futures, IIED, London.
- Pimentel, D., 2007. Area-wide pest management: environmental, economic and food issues. In *Area wide control for insect pests*. The Netherlands: IAEA, Springer, pp. 35–47.
- Pingali, P.L., 2012. Green Revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31), pp.12302–12308. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22826253> [Accessed July 31, 2012].
- Points, I., 2011. <http://influentialpoints.com/index.htm>. *Stats with attitude*.
- Polajnar, J. et al., 2015. Manipulating behaviour with substrate-borne vibrations - potential for insect pest control. *Pest management science*, 71(1), pp.15–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24962656> [Accessed December 27, 2014].
- Popodi, E.M. et al., 2008. Cooperativity between the beta-tubulin carboxy tail and the body of the molecule is required for microtubule function. *Cell Motility and the Cytoskeleton*, 65(12), pp.955–963.
- Popodi, E.M. et al., 2005. The proximal region of the beta-tubulin C-terminal tail is sufficient for axoneme assembly. *Cell motility and the cytoskeleton*, 62(1), pp.48–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16080206> [Accessed October 4, 2012].

- Posnien, N. et al., 2009. RNAi in the red flour beetle (*Tribolium*). *Cold Spring Harbor protocols*, 2009(8), p.pdb.prot5256. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20147232> [Accessed January 3, 2012].
- Pretty, J., 2008. Agricultural sustainability: concepts, principles and evidence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(July 2007), pp.447–465.
- Price, D.R.G. & Gatehouse, J.A., 2008. RNAi-mediated crop protection against insects. *Trends in Biotechnology*, 26(May), pp.393–400.
- Qaim, M. & Kouwer, S., 2013. Genetically Modified Crops and Food Security M. L. Alvarez, ed. *PLoS ONE*, 8(6), p.e64879. Available at: <http://dx.plos.org/10.1371/journal.pone.0064879> [Accessed June 5, 2013].
- Raff, E.C., 1997. Microtubule Architecture Specified by a beta-Tubulin Isoform. *Science*, 275(5296), pp.70–73. Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.275.5296.70> [Accessed May 12, 2014].
- Raff, E.C. et al., 1982. Regulation of tubulin gene expression during embryogenesis in *Drosophila melanogaster*. *Cell*, 28(1), pp.33–40. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6802501>.
- Raphael, K. a et al., 2011. Germ-line transformation of the Queensland fruit fly, *Bactrocera tryoni*, using a piggyBac vector in the presence of endogenous piggyBac elements. *Genetica*, 139(1), pp.91–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20859652> [Accessed March 1, 2013].
- Rathke, C. et al., 2014. Chromatin dynamics during spermiogenesis. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1839(3), pp.155–168. Available at: <http://dx.doi.org/10.1016/j.bbagr.2013.08.004>.
- Ravi Ram, K. & Wolfner, M.F., 2007. Seminal influences: *Drosophila* Acps and the molecular interplay between males and females during reproduction. *Integrative and comparative biology*, 47(3), pp.427–45. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21672851> [Accessed September 11, 2013].
- Reade, M.C. et al., 2008. Bench-to-bedside review: avoiding pitfalls in critical care meta-analysis-funnel plots, risk estimates, types of heterogeneity, baseline risk and the ecologic fallacy. *Critical care (London, England)*, 12(4), p.220. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2575558/> [tool=pmcentrez&rendertype=abstract] [Accessed November 24, 2013].
- Reeves, R.G. et al., 2012. Scientific standards and the regulation of genetically modified insects. *PLoS neglected tropical diseases*, 6(1), p.e1502. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3269433/> [tool=pmcentrez&rendertype=abstract] [Accessed March 14, 2012].
- Reimer, B.S., 2012. Nuclear Migration and Spermatid Elongation During Spermiogenesis in *Drosophila Pseudoobscura*. Seton Hall University.
- Rendon, P. et al., 2006. The effects of sterile males and two braconid parasitoids, *Fopius arisanus* (Sonan) and *Diachasmimorpha krausii* (Fullaway) (Hymenoptera), on caged populations of Mediterranean fruit flies, *Ceratitis capitata* (Wied.) (Diptera: Tephritidae) at various sites. *Biological Control*, 36(2), pp.224–231.

- Richards, S. et al., 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature*, 452(7190), pp.949–55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18362917> [Accessed June 10, 2011].
- Rietema, K., 1991. *The Dynamics of Fine Powders*, Barking, Essex: Elsevier Science Publishers Ltd.
- Ritz, C. & Streibig, J.C., 2005. Bioassay Analysis using R. *Journal of Statistical Software*, 12(5). Available at: <http://www.bioassay.dk>.
- Robertson, S.L., Cushing, J.M. & Costantino, R.F., 2012. Life stages: interactions and spatial patterns. *Bulletin of mathematical biology*, 74(2), pp.491–508. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22135094> [Accessed April 11, 2012].
- Robinson, A.S., 2005. Genetic Basis of the Sterile Insect Technique. In V. a. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 95–114.
- Robinson, A.S., Franz, G. & Atkinson, P.W., 2004. Insect transgenesis and its potential role in agriculture and human health. *Insect biochemistry and molecular biology*, 34(2), pp.113–20. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14871607> [Accessed July 24, 2012].
- Ronald, P., 2011. Plant genetics, sustainable agriculture and global food security. *Genetics*, 188(1), pp.11–20. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3120150/> [tool=pmcentrez&rendertype=abstract] [Accessed July 1, 2011].
- Rose, a B. & Beliakoff, J. a, 2000. Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. *Plant physiology*, 122(2), pp.535–542.
- Rubin, G.M. & Spradling, a C., 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science (New York, N.Y.)*, 218(4570), pp.348–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6289436>.
- Salinas, P., 1985. Studies on Diamondback moth in Venezuela with reference to other Latin american countries. In *Diamondback Moth Management*. The Corporate Storyteller, Sydney, Australia, pp. 17–24.
- Salz, H.K., 2011. Sex determination in insects: a binary decision based on alternative splicing. *Current opinion in genetics & development*, 21(4), pp.395–400. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3134629/> [tool=pmcentrez&rendertype=abstract] [Accessed August 1, 2011].
- Sanchez, L. & Sánchez, L., 2008. Sex-determining mechanisms in insects. *The International journal of developmental biology*, 52(7), pp.837–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18956315> [Accessed September 5, 2011].
- Santel, a et al., 2000. The initiator element of the *Drosophila* beta2 tubulin gene core promoter contributes to gene expression in vivo but is not required for male germ-cell specific expression. *Nucleic acids research*, 28(6), pp.1439–46. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC111050/> [tool=pmcentrez&rendertype=abstract]
- Sarkar, A. et al., 2003. Molecular evolutionary analysis of the widespread piggyBac transposon family and related “domesticated” sequences. *Molecular genetics and genomics MGG*, 270, pp.173–180.

- Sassone-Corsi, P., 2002. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science (New York, N.Y.)*, 296(5576), pp.2176–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12077401>.
- Saussede-aim, J., Dumontet, C. & Lyon, U., 2009. Regulation of tubulin expression : Multiple overlapping mechanisms. *International Journal of Medicine and Medical Sciences*, 1(8), pp.290–296.
- Savary, S. et al., 2006. Quantification and modeling of crop losses: a review of purposes. *Annual review of phytopathology*, 44, pp.89–112.
- Sayer, J. & Cassman, K., 2013. Agricultural Innovation to Protect the Environment. *Pnas*, 110(21), pp.8345–8348. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3666751&tool=pmcentrez&rendertype=abstract>.
- Schellhorn, N. et al., 2004. The use of dyes to mark populations of beneficial insects in the field. *International Journal of Pest Management*, 50(3), pp.153–159. Available at: <http://www.tandfonline.com/doi/abs/10.1080/09670870410001731862> [Accessed September 15, 2011].
- Schetelig, M.F. et al., 2012. Male only progeny in *Anastrepha suspensa* by RNAi-induced sex reversion of chromosomal females. *Insect Biochemistry and Molecular Biology*, 42(1), pp.51–57. Available at: <http://dx.doi.org/10.1016/j.ibmb.2011.10.007>.
- Schetelig, M.F. & Handler, A.M., 2012. A transgenic embryonic sexing system for *Anastrepha suspensa* (Diptera: Tephritidae). *Insect biochemistry and molecular biology*, 42(10), pp.790–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22858603> [Accessed April 25, 2013].
- Schetelig, M.F. & Handler, A.M., 2013. Germline transformation of the spotted wing drosophilid, *Drosophila suzukii*, with a piggyBac transposon vector. *Genetica*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23564446> [Accessed April 11, 2013].
- Schinko, J.B. et al., 2010. Functionality of the GAL4/UAS system in *Tribolium* requires the use of endogenous core promoters. *BMC developmental biology*, 10, p.53. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2882914&tool=pmcentrez&rendertype=abstract>.
- Schliekelman, P. & Gould, F., 2000a. Pest control by the introduction of a conditional lethal trait on multiple loci: potential, limitations, and optimal strategies. *Journal of economic entomology*, 93(6), pp.1543–1565.
- Schliekelman, P. & Gould, F., 2000b. Pest control by the release of insects carrying a female-killing allele on multiple loci. *Journal of economic entomology*, 93(6), pp.1566–1579.
- Schnell, J. et al., 2014. A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. *Transgenic Research*, 24, pp.1–17. Available at: <http://link.springer.com/10.1007/s11248-014-9843-7>.
- Schnell, J.D. & Hicke, L., 2003. Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *The Journal of biological chemistry*, 278(38), pp.35857–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12860974> [Accessed August 18, 2014].
- Scolari, F. et al., 2008. Fluorescent sperm marking to improve the fight against the pest insect *Ceratitis capitata* (Wiedemann; Diptera: Tephritidae). *New biotechnology*, 25(1), pp.76–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18504022> [Accessed October 29, 2011].

- Scolari, F. et al., 2011. Safe and fit genetically modified insects for pest control: from lab to field applications. *Genetica*, 139(1), pp.41–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20725766> [Accessed October 31, 2011].
- Scott, J.G. et al., 2013. Towards the elements of successful insect RNAi. *Journal of insect physiology*, 59(12), pp.1212–21. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24041495> [Accessed October 17, 2014].
- Scott, M.J., Heinrich, J.C. & Li, X., 2004. Progress towards the development of a transgenic strain of the Australian sheep blowfly (*Lucilia cuprina*) suitable for a male-only sterile release program. *Insect biochemistry and molecular biology*, 34(2), pp.185–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14871615> [Accessed April 25, 2013].
- Seirin Lee, S. et al., 2013. Modelling Aedes aegypti mosquito control via transgenic and sterile insect techniques: Endemics and emerging outbreaks. *Journal of Theoretical Biology*, 331, pp.78–90. Available at: <http://dx.doi.org/10.1016/j.jtbi.2013.04.014>.
- Servedio, M.R. et al., 2014. Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. *PLoS Biology*, 12(12), p.e1002017. Available at: <http://dx.plos.org/10.1371/journal.pbio.1002017> [Accessed December 9, 2014].
- Sethuraman, N. et al., 2007. Post-integration stability of piggyBac in Aedes aegypti. *Insect Biochemistry and Molecular Biology*, 37, pp.941–951.
- Shagin, D. a et al., 2004. GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Molecular biology and evolution*, 21(5), pp.841–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14963095> [Accessed March 4, 2012].
- Shinmyo, Y. et al., 2004. piggyBac -mediated somatic transformation of the two-spotted cricket , Gryllus bimaculatus. *Development Growth and Differentiation*, 46, pp.343–349.
- Shukla, J.N. & Palli, S.R., 2012a. Doublesex target genes in the red flour beetle, *Tribolium castaneum*. *Scientific reports*, 2, p.948. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3517981&tool=pmcentrez&rendertype=abstract> [Accessed December 21, 2012].
- Shukla, J.N. & Palli, S.R., 2014. Production of all female progeny: an evidence for the presence of male sex determination factor on Y chromosome. *The Journal of Experimental Biology*, 217(February), pp.1653–1655. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24577442>.
- Shukla, J.N. & Palli, S.R., 2012b. Sex determination in beetles: Production of all male progeny by Parental RNAi knockdown of transformer. *Scientific Reports*, 2, pp.1–9. Available at: <http://www.nature.com/doifinder/10.1038/srep00602> [Accessed August 27, 2012].
- Shukla, J.N. & Palli, S.R., 2013. *Tribolium castaneum* Transformer-2 regulates sex determination and development in both males and females. *Insect biochemistry and molecular biology*, 43(12), pp.1125–1132. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24056158> [Accessed November 13, 2013].
- Siebert, K.S. et al., 2008. Tubulin superfamily genes in *Tribolium castaneum* and the use of a Tubulin promoter to drive transgene expression. *Insect biochemistry and molecular biology*, 38(8), pp.749–55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18625397> [Accessed September 12, 2011].
- Silverman, B.W., 2002. Density estimation for statistics and data analysis. In *Monographs on Statistics and Applied Probability*. London: Chapman & Hall, pp. 1–22.

- Sim, H. et al., 2011. Acupuncture for carpal tunnel syndrome: a systematic review of randomized controlled trials. *The journal of pain : official journal of the American Pain Society*, 12(3), pp.307-14. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21093382> [Accessed September 27, 2013].
- Simmons, G.S. et al., 2011. Field performance of a genetically engineered strain of pink bollworm. *PloS one*, 6(9), p.e24110. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3172240/> [Accessed October 21, 2011].
- Simmons, G.S. et al., 2010. Improved quality management to enhance the efficacy of the sterile insect technique for lepidopteran pests. *Journal of Applied Entomology*, 134(3), pp.261–273. Available at: <http://doi.wiley.com/10.1111/j.1439-0418.2009.01438.x> [Accessed October 8, 2012].
- Singh, A.D. et al., 2013. Oral delivery of double-stranded RNA in larvae of the yellow fever mosquito, *Aedes aegypti*: implications for pest mosquito control. *Journal of insect science (Online)*, 13, p.69. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835047/> [Accessed October 21, 2011].
- Smith, A.F.M. & Gelfand, A.E., 1992. Bayesian Statistics Without Tears : A Sampling-Resampling Perspective. *The American Statistician*, 46(2), pp.84–88.
- Smith, R.C. et al., 2007. Testis-specific expression of the beta2 tubulin promoter of *Aedes aegypti* and its application as a genetic sex-separation marker. *Insect molecular biology*, 16(1), pp.61–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17257209> [Accessed October 29, 2011].
- Sokoloff, A., 1972. The Biology of *Tribolium* with special emphasis on genetic aspects,. Clarendon Press.
- Soopaya, R. et al., 2011. Radiation Biology and Inherited Sterility of Light Brown Apple Moth (Lepidoptera : Tortricidae): Developing a Sterile Insect Release Program Radiation Biology and Inherited Sterility of Light Brown Apple Moth (Lepidoptera : Tortricidae): Developing a. *Biological Control*, 104(6), pp.1999–2008.
- Sørensen, J.G., Addison, M.F. & Terblanche, J.S., 2012. Mass-rearing of insects for pest management: Challenges, synergies and advances from evolutionary physiology. *Crop Protection*, 38, pp.87–94. Available at: <http://dx.doi.org/10.1016/j.cropro.2012.03.023>.
- Sotiropoulos, V. & Kaznessis, Y.N., 2007. Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes. *BMC systems biology*, 1, p.7.
- Steger, K., 2001. Haploid spermatids exhibit translationally repressed mRNAs. *Anatomy and Embryology*, 203(5), pp.323–334.
- Subramanian, S., Rabindra, R.J. & Sathiah, N., 2010. Economic threshold for the management of *Plutella xylostella* with granulovirus in cauliflower ecosystem. *Phytoparasitica*, 38(1), pp.5–17.
- Suckling, D.M. et al., 2014. From integrated pest management to integrated pest eradication: technologies and future needs. *Pest management science*, 70(2), pp.179–89. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24155254> [Accessed November 21, 2014].
- Sumitani, M. et al., 2003. Germline transformation of the sawfly, *Athalia rosae* (Hymenoptera: Symphyta), mediated by a piggyBac-derived vector. *Insect Biochemistry and Molecular Biology*,

- 33(4), pp.449–458. Available at:
<http://linkinghub.elsevier.com/retrieve/pii/S0965174803000092> [Accessed March 1, 2013].
- Suzuki, M.G., 2010. Sex determination: insights from the silkworm. *Journal of genetics*, 89(3), pp.357–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22290220>.
- Swallow, J.G. & Wilkinson, G.S., 2002. The long and short of sperm polymorphisms in insects. *Biological reviews of the Cambridge Philosophical Society*, 77(2), pp.153–82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12056745>.
- Takahashi, J. et al., 2008. Interplasmid transposition of the piggyBac element in embryos of the honey bee, *Apis mellifera*. *Journal Of Apicultural Research*, 47, pp.304–309.
- Taleb, N.N., 2007. Black Swans and the Domains of Statistics. *The American Statistician*, 61(3), pp.198–200. Available at: <http://www.tandfonline.com/doi/abs/10.1198/000313007X219996> [Accessed March 7, 2013].
- Talekar, N. & Shelton, A., 1993. Biology, Ecology and Management of the Diamondback Moth. *Annual Reviews of Entomology*, 38(92), pp.275–301.
- Tamura, T. et al., 2000. Germline transformation of the silkworm *Bombyx mori* L. using a piggyBac transposon-derived vector. *Nature biotechnology*, 18(1), pp.81–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10625397>.
- Tan, a. et al., 2013. Transgene-based, female-specific lethality system for genetic sexing of the silkworm, *Bombyx mori*. *Proceedings of the National Academy of Sciences*, pp.1–5. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1221700110> [Accessed April 9, 2013].
- Tanaka, T. & Tagashira, E., 1998. Parasitic castration of *Pseudaletia separata* by *Cotesia kariyai* and its association with polydnavirus gene expression. *Journal of insect physiology*, 44(9), pp.733–744. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12769869>.
- Terenius, O. et al., 2011. RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology*, 57(2), pp.231–245.
- Thailayil, J. et al., 2011. Spermless males elicit large-scale female responses to mating in the malaria mosquito *Anopheles gambiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(33), pp.13677–81. Available at:
<http://www.ncbi.nlm.nih.gov/article/fcgi?artid=3158155&tool=pmcentrez&rendertype=abstract> [Accessed August 7, 2013].
- Theilmann, D. a & Stewart, S., 1992. Molecular analysis of the trans-activating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology*, 187(1), pp.84–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1736546>.
- Thomas, D.D., 2000. Insect Population Control Using a Dominant, Repressible, Lethal Genetic System. *Science*, 287(5462), pp.2474–2476. Available at:
<http://www.science.org/cgi/doi/10.1126/science.287.5462.2474> [Accessed July 13, 2011].
- Thompson, J.R., 2001. The Age of Tukey. *Technometrics*, 43(3), pp.256–265.
- Thresher, R. et al., 2009. Development of repressible sterility to prevent the establishment of feral populations of exotic and genetically modified animals. *Aquaculture*, 290(1-2), pp.104–109. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0044848609001574> [Accessed June 21, 2013].

- Tomoyasu, Y. et al., 2008. Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in *Tribolium*. *Genome biology*, 9(1), p.R10.
- Traut, W. et al., 2006. Phylogeny of the sex-determining gene Sex-lethal in insects. *Genome*, 49, pp.254–262.
- Tscharntke, T. et al., 2012. Global food security, biodiversity conservation and the future of agricultural intensification. *Biological Conservation*, 151(1), pp.53–59. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0006320712000821> [Accessed October 4, 2012].
- Tsubota, T. et al., 2014. Identification of a novel strong and ubiquitous promoter/enhancer in the silkworm *Bombyx mori*. *G3*, 4(7), pp.1347–57. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24875626> [Accessed January 6, 2015].
- Tukey, J., 1977. *Exploratory Data Analysis*, Addison-Wesley.
- Vandenbroucke, J.P., 2008. Observational research, randomised trials, and two views of medical science. *PLoS medicine*, 5(3), p.e67. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2265762&tool=pmcentrez&rendertype=abstract> [Accessed April 2, 2013].
- Varshavsky, A., 2005. Ubiquitin Fusion Techniques and Related Methods. *Methods in Enzymology*, 399, p.777.
- Verkhusha, V. V & Lukyanov, K. a, 2004. The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nature biotechnology*, 22(3), pp.289–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14990950> [Accessed June 14, 2011].
- Viktorinová, I. & Wimmer, E. a, 2007. Comparative analysis of binary expression systems for directed gene expression in transgenic insects. *Insect biochemistry and molecular biology*, 37(3), pp.246–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17296499> [Accessed March 15, 2012].
- Visser, I., 2011. Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *Journal of Mathematical Psychology*, 55(6), pp.403–415. Available at: <http://dx.doi.org/10.1016/j.jmp.2011.08.002>.
- Voigt, C., 2011. PROGRAMMING CELLS: TOWARDS AN AUTOMATED “GENETIC COMPILER”. In Institute of Medicine (US) Forum on Microbial Threats. The Science and Applications of Synthetic and Systems Biology: Workshop Summary. National Academies Press, Washington. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK84438/>.
- Vreysen, M.J.B., Carpenter, J.E. & Marec, F., 2010. Improvement of the sterile insect technique for codling moth *Cydia pomonella* (Linnaeus) (Lepidoptera Tortricidae) to facilitate expansion of field application. *Journal of Applied Entomology*, 134(3), pp.165–181. Available at: <http://doi.wiley.com/10.1111/j.1439-0418.2009.01430.x> [Accessed October 8, 2012].
- Waage, J.K. & Mumford, J.D., 2008. Agricultural biosecurity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1492), pp.863–76. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2610114&tool=pmcentrez&rendertype=abstract> [Accessed July 19, 2011].
- Wade, M. & Stevens, L., 1985. Microorganism Mediated Reproductive Isolation in Flour Beetles (Genus *Tribolium*). *Science*, 227(4686), pp.527–528.
- Walton, J.R. & Marshall, J.T., 1987. Zygote Viability in Gene Transfer. *Biology of reproduction*, 37, pp.957–967.

- Wang, J. et al., 2008. Large diversity of the piggyBac-like elements in the genome of *Tribolium castaneum*. *Insect Biochemistry and Molecular Biology*, 38(4), pp.490–498. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18342253>.
- Wang, L. & Xu, C., 2014. Role of microRNAs in mammalian spermatogenesis and testicular germ cell tumors. *Reproduction*, 149(3), pp.R127–R137. Available at: <http://www.reproduction-online.org/cgi/doi/10.1530/REP-14-0239>.
- Warren, I. a, Fowler, K. & Smith, H., 2010. Germline transformation of the stalk-eyed fly, *Teleopsis dalmanni*. *BMC molecular biology*, 11(1), p.86. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2999598&tool=pmcentrez&rendertype=abstract> [Accessed May 2, 2013].
- Wasserman, L., 2006. *All of nonparametric statistics* G. Casella, ed., New York: Springer Science.
- Wells, D.J., 2004. Gene therapy progress and prospects: electroporation and other physical methods. *Gene therapy*, 11(18), pp.1363–1369.
- White-Cooper, H., 2010. Molecular mechanisms of gene regulation during *Drosophila* spermatogenesis. *Reproduction (Cambridge, England)*, 139(1), pp.11–21. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19755484> [Accessed June 9, 2013].
- White-Cooper, H. & Bausek, N., 2010. Evolution and spermatogenesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1546), pp.1465–80. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2871925&tool=pmcentrez&rendertype=abstract> [Accessed June 10, 2013].
- White-Cooper, H. & Davidson, I., 2011. Unique aspects of transcription regulation in male germ cells. *Cold Spring Harbor perspectives in biology*, 3(7), pp.1–16. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3119912&tool=pmcentrez&rendertype=abstract>.
- Whitten, M. & Mahon, R., 2005. Misconceptions and constraints. In V. a. Dyck, J. Hendrichs, & A. S. Robinson, eds. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer, pp. 601–626.
- Whyard, S. et al., 2015. Silencing the buzz: a new approach to population suppression of mosquitoes by feeding larvae double-stranded RNAs. *Parasites & Vectors*, 8(1), pp.1–11. Available at: <http://www.parasitesandvectors.com/content/8/1/96>.
- Whyard, S., Singh, A.D. & Wong, S., 2009. Ingested double-stranded RNAs can act as species-specific insecticides. *Insect biochemistry and molecular biology*, 39(11), pp.824–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19815067> [Accessed March 3, 2013].
- Wickham, H., 2014. Tidy Data. *Journal of Statistical Science*, 59(10). Available at: <http://courses.had.co.nz.s3-website-us-east-1.amazonaws.com/12-rice-bdsi/slides/07-tidy-data.pdf> [Accessed October 1, 2014].
- Wickham, H. & Stryjewski, L., 2010. 40 years of boxplots. *Under review*, pp.1–17. Available at: <http://had.co.nz/stat645/project-03/boxplots.pdf> [Accessed October 1, 2014].
- Wilson, K. & Hardy, I.C.W., 2002. Statistical analysis of sex ratios: an introduction. In *Sex ratios: concepts and research methods*. pp. 48–92. Available at: http://books.google.com/books?hl=en&lr=&id=Pj7qnU-cF6YC&oi=fnd&pg=PA48&dq=Statistical+analysis+of+sex+ratios:+an+introduction&ots=H8q4_3Ws2Y&sig=l_elofubi0PDcw70dGoudalQfhY.

- Winkler, R., Smith, J. & Fryback, D., 2002. The Role of Informative Priors in Zero-Numerator Problems : Being Conservative Versus Being Candid. *Statistical Practice*, 56(1), pp.1–4.
- Woltjen, K. et al., 2011. Transgene-Free Production of Pluripotent Stem Cells Using piggyBac Transposons. *Methods In Molecular Biology Clifton Nj*, 767, pp.87–103.
- Wright, J. a et al., 2013. IPB7 transposase behavior in *Drosophila melanogaster* and *Aedes aegypti*. *Insect biochemistry and molecular biology*, 43(10), pp.899–906. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23835045> [Accessed November 13, 2013].
- Wu, F. & Butz, W., 2004. The future of genetically modified crops: lessons from the Green Revolution. R. Cooperation, ed., Santa Monica.
- Wu, X. & Burgess, S.M., 2004. Integration target site selection for retroviruses and transposable elements. *Cellular and molecular life sciences : CMLS*, 61(19-20), pp.2588–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15526164> [Accessed April 2, 2013].
- Xu, B., Feng, X. & Burdine, R.D., 2010. Categorical data analysis in experimental biology. *Developmental biology*, 348(1), pp.3–11. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2900000/> [Accessed July 19, 2012].
- Xu, J. et al., 2014. Transgenic characterization of two testis-specific promoters in the silkworm, *Bombyx mori*. *Insect molecular biology*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25387604> [Accessed December 2, 2014].
- Yi, H. et al., 2014. Effects of global warming on mosquitoes & mosquito-borne diseases and the new strategies for mosquito control. *Entomological Research*, 44(6), pp.215–235. Available at: <Go to ISI>://WOS:000345459900001.
- Yu, N. et al., 2013. Delivery of dsRNA for RNAi in insects: an overview and future directions. *Insect Science*, 20(1), pp.4–14. Available at: <http://doi.wiley.com/10.1111/j.1744-7917.2012.01534.x> [Accessed March 15, 2013].
- Yusa, K. et al., 2011. A hyperactive piggyBac transposase for mammalian applications. *Proceedings of the National Academy of Sciences of the United States of America*, 108, pp.1531–1536.
- Zhang, X., 2015. Chitosan/interfering RNA nanoparticle mediated gene silencing in disease vector mosquito larvae. *Journal of visualized experiments*, (97).
- Zhao, J. et al., 2010. FlyTED: the *Drosophila* Testis Gene Expression Database. *Nucleic acids research*, 38(Database issue), pp.D710–5. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2808924/> [Accessed March 21, 2014].
- Zhuang, L. et al., 2010. The relationship between internal domain sequences of piggyBac and its transposition efficiency in BmN cells and *Bombyx mori*. *Acta Biochim Biophys*, 42(6), pp.426–431.
- Zimowska, G.J. & Handler, A.M., 2006. Highly conserved piggyBac elements in noctuid species of Lepidoptera. *Insect Biochemistry and Molecular Biology*, 36, pp.421–428.
- Zimowska, G.J., Nirmala, X. & Handler, A.M., 2009. The beta2-tubulin gene from three tephritid fruit fly species and use of its promoter for sperm marking. *Insect biochemistry and molecular biology*, 39(8), pp.508–15. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19520163> [Accessed October 29, 2011].

