Horticulture Innovation Australia

Final Report

Using Classical Genetics and Epigenetics to Make SIT Flies Fitter

Anthony Clarke
Queensland University of Technology

Project Number: AI14002

AI14002

This project has been funded by Horticulture Innovation Australia Limited with co-investment from the Queensland University of Technology and funds from the Australian Government.

ISBN 0 7341 3898 9

Published and distributed by:
Horticulture Innovation Australia Limited
Level 8, 1 Chifley Square
Sydney NSW 2000
Tel: (02) 8295 2300
Fax: (02) 8295 2399

Contents

Summary

Background: The Sterile Insect Technique (SIT) is an effective and eco-friendly approach to combat pest tephritid fruit flies. However, a key limitation in the SIT can be poor quality sterile males and their inability to compete with wild males for females. Feeding on the so called 'male lures' (= plant derived phenylpropanoids long used in fruit fly surveillance and monitoring) generally improve a male's mating success and so could be used to enhance the quality of SIT males. However, to feed lure to SIT males would need them to be held to sexual maturity, as opposed to being released as pupae or very young adults, and this would add significantly to the cost of the SIT. Hence, an approach that provides the benefits of lure exposure, without holding adults, is sought-after as it could greatly advantage the SIT with minimal costs.

The effects of male lures reported in *Bactrocera* fruit flies are complex. For instance, the male lure zingerone changes the genetic expression and energy metabolism of the Queensland fruit fly, *Bactrocera tryoni* (Qfly), rendering males physically fitter and more competitive in mating. Further, Qfly offspring sired by lure-fed fathers were found to forage for lures better, and this suggests possible epigenetic modifications in males after lure feeding that is being passed on to the offspring through females. The known changes in genetic expression, and the hypothesised changes in the epigenome, strongly infer that the lure 'factor' is modifying the expression of the underlying genotype that regulates fitness in males. If we are able to screen the genetic makeup that is being influenced by the lure, then it is theoretically possible to manipulate the genome of a SIT selected male line such that they express the benefits of lure exposure, without the need for the actual exposure.

Objectives: The first objective of this transformational postdoctoral fellowship was to mentor a postdoctoral researcher to becoming an independent scientist in the field of advanced fruit fly genetics.

The second objective was to develop a deeper understanding on the genetics of *Bactrocera tryoni* (=Qfly) male fitness to help to improve the quality of the sterile males. To understand the genetic and epigenetic changes occurring in Qfly after lure feeding, we used classical genetics, genotyping-by-sequencing and epigenetics approaches. Using these techniques we tested: (i) heritability of the lure foraging trait; (ii) histone modifications in males after lure feeding; and (iii) gene expression changes in females mated with the lure-fed males. The epigenetics research, especially, is world leading and truly transformational not just in fruit flies, but in any field of applied entomology.

Key outputs: The output of Objective 1 is a qualified researcher with substantial experience in high throughput NGS (next generation sequencing) technologies, chromatin immunoprecipitation (to study epigenetics), restriction site associated DNA polymorphism (to study heritability of behavioural traits) and transcriptomics (to understand gene expression). Training in these novel techniques will assist in the study and understanding of the genetic mechanisms mediating fitness in any insect.

The outputs of Objective 2 are three research manuscripts directly related to SIT (two preliminary drafts appended), one manuscript on the comparative ecology of Qfly and its sibling species lesser Qfly (*Bactrocera neohumeralis*), and recommendations for strategic future research. The first manuscript presents a methodology to study epigenetic changes in tephritid fruit flies. The second manuscript demonstrates genes regulating mating and re-mating in female Qfly. The manuscript on *B. neohumeralis* presents host and habitat use patterns in *B. tryoni* and *B. neohumeralis* that demonstrates no niche segregation between these two species. A fourth manuscript on heritability of lure foraging trait and epigenetic histone modifications after lure feeding is at the discussion and analysis phase and hence the draft is not appended. Manuscripts have not submitted during the life of the project simply because of

the time it took to generate and analyse the very large amount of genomic data gathered. With most manuscripts in advanced stages of production, we anticipate all publications submitted by the end of the 2016/2017 financial year.

In addition to the manuscripts, we have generated a vast amount of quality genetic data that will be available to other researchers. These resources will be highly useful for further targeted research in understanding and manipulating the genes and gene families mediating fitness in male (and female) Qfly.

Outcomes of the project: By generating quality genetic resources, and by studying genetic and epigenetic changes in Qfly, this project made significant advancement towards understanding the genetic mechanisms regulating mating and fitness traits in Qfly. The genetic data generated from this project will add to the resources to mediate further research to develop high quality sterile males for the successful application of the SIT for *B. tryoni* control. In addition, this project aided in the advanced education and mentoring of a now highly qualified research scientist with an exceptional capacity to significantly contribute to the application of genomics and epigenomics to applied pest management in horticulture and other agricultural sectors.

Recommendations: This study advances the knowledge of how male lures may be used in SIT programs against *Bactrocera* pest species, additional to that already known via aromatherapy techniques (i.e. direct exposure of mass-reared adult flies to lures). The studies on heritability of the lure foraging trait confirmed earlier published work that offspring of lure-fed males are more effective in finding the lure source. The heritable nature of this effect is demonstrated in this study for the first time, and hints that mass culturing facilities can possibly utilize male lures as a dietary supplement in their permanent breeding lines which would, increasingly, positively select for the male lure search response and so increase the fitness of released flies. However, further large scale, dedicated empirical studies are needed to validate these results before adding lures as a diet supplement in rearing lines.

The genetic data we developed can be used to further understand the fitness of males so that we can produce high quality 'fitter' sterile lines. While the data from the genotyping-by-sequencing study identify potential SNPs after lure feeding, the data from histone modification experiments identify genes and gene families mediating male fitness. In addition, we have generated transcriptome data specific to female *B. tryoni* for the first time, and this will benefit future studies understanding the genetic mechanisms mediating female re-mating.

Keywords

SIT, Queensland fruit fly, *Bactrocera tryoni*, mating, reproduction, fitness, Tephritidae, phytochemicals, lures, mating success, attractants, epigenetics, foraging, gene expression, transcriptomics, histone modifications, cue lure, methyl eugenol, zingerone, *Bactrocera neohumeralis*.

Introduction

The Queensland fruit fly (Qfly), *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae), is a major insect pest of national significance and an ongoing threat to horticultural industries (Clarke et al. 2011). Qfly infests a wide range of horticultural crops, and monetary losses due to monitoring, surveillance, management and lost market access were estimated at $28.5 million per annum in 2000, and over $128 million for the years 2003-2008 (PHA 2008). Management of Qfly is mainly through bait trapping, attract-and-kill, male annihilation and the Sterile Insect Technique (SIT). While these approaches, when combined with insecticide cover sprays, have proved successful in controlling the fly, the regulatory loss of dimethoate and fenthion means that the controls need to be refined so that they become more effective.

Of the non-pesticide controls, the SIT is considered the one with the greatest potential, as it can be applied as a wide area approach to control source fruit fly populations. SIT has been practiced in many parts of the country, and is recognized as an environmentally benign approach for management of fruit fly pests (Dyck et al. 2005). However, a key limitation can be the quality of sterile males and their ability to out-compete wild males for female mates (Meats et al. 2004; Weldon, 2005). Sterilization, coupled with mass rearing, induce strong selection pressures (Henning et al. 2000; Calkins & Parker 2005) and may ultimately lead to poor quality males with reduced flight capacity, greater bouts of inactivity leading to an inability to forage for natural resources, and poor courtship songs (Weldon et al. 2010; Mankin et al., 2008). Such issues can directly compromise the success of the SIT.

Within the fruit fly genera *Bactrocera* and *Zeugodacus*, feeding on plant secondary phenylpropanoids, chemicals known commonly in the fruit fly literature as 'male lures', generally improves a male fly's mating success. The benefits of exposure to other classes of plant secondary chemicals are known for other pest fruit flies (such as the Mediterranean fruit fly) and is incorporated into SIT programs by holding adult flies and exposing them to the chemicals, a practice known as aromatherapy (Shelly et al. 2004). However, it adds significantly to the cost of SIT as flies need to be held for several days in the adult stage, as opposed to being released as pupae or newly emerged adults. Hence, an approach that provides the benefits of lure exposure, without holding adults, should be a sought-after technique that would greatly benefit SIT.

In earlier work, we found that male lures changed the expression of energy metabolism genes and ultimately the physiology of male *B. tryoni* (Kumaran et al. 2013; 2014). Further, in a unique study, we found the male offspring sired by *B. tryoni* males fed on phytochemical lures had greater foraging ability for the lures, which would directly enhance their fitness (*sensu* runaway sexual selection, Kumaran & Clarke 2014). This lure effect is indirect in offspring from the father through the mother, and strongly implies an epigenetic effect. Epigenetic effects are heritable changes in gene expression that do not involve changes in underlying DNA sequence, i.e. a cross-generational change in phenotype, but not genotype. We believe this is a key to helping make males fitter using genetic approaches and could be used in the SIT so that released males could compete with wild males more effectively. We suspect that the lure factor is modifying expression of the underlying genotype that governs reproductive fitness in male fruit flies. If we are able to screen the genetic makeup that is being influenced by the lure, then it is possible to manipulate the genome of an SIT selected male line such that they express the benefits of lure exposure, without the need for actual exposure (as practiced in Medfly).

Given this background, this project was funded to further understand the genetic changes in Qfly after lure feeding and to test the heritability of the lure foraging trait. We initially proposed to HIA to use 'simple' transcriptome based approaches to identify gene expression in the offspring of lure-fed males to understand the genetic changes occurring. However, with collaboration from geneticists and additional in-kind support from QUT, we used the high throughput technologies ChIP-seq (chromatin immuno-precipitation and sequencing) and RAD-seq (Restriction site associated DNA polymorphism and

sequencing) approaches instead of a transcriptome approach. This alternative research methodology was outlined and accepted in milestone reports 104 and 105. In line with the original project aims, we also studied the heritability of lure foraging in Qfly using classical genetic trials for nine generations in large field enclosures. Using the parental and lure selected lines, we performed genotyping by sequencing using the RAD-seq approach to obtain genetic resources to identify single nucleotide polymorphisms (SNPs) in selected lines. The results obtained from the heritability study revealed that the lure foraging trait is heritable in Qfly and we can possibly select for populations that can more effectively forage for male lures. From the RAD-seq approach we sequenced parental and F9 selected and normal unselected lines and generated 49.12 gb of sequence data with 327,493,943 quality reads that will help identify potential SNPs in lure selected populations.

Additional to the original project goals (of documenting genetic changes following classical selection), we studied histone modifications after lure feeding in Qfly males to understand the epigenetic mechanism regulating male fitness; and transcriptional changes in females after mating with lure-fed males to understand the genetic mechanisms mediating mating and re-mating in females. Both studies significantly value added to the project and genetic resources generated. Histone modifications help us identify the potential epigenetic mechanisms regulating male fitness, while the transcriptional changes help us understand the mechanisms driving females to re-mate. The ChIP-seq study revealed histone modifications in Qfly for the first time, and explicitly the changes found in flies fed with lures: we identified ~200 peaks for H3K4me3 and ~160 peaks for H3K36me3 antibodies differentially expressed in lure-fed males when compared with normal males through Diffbind analyses. The female transcriptome study revealed 89 genes differentially expressed in females when mated with lure-fed males compared to females mated with normal males, with 70 genes up-regulated and 19 genes down regulated. Investigation of gene categories revealed significant enrichment of several GO terms. Together, these studies provide novel resources to understanding the genetic basis of foraging, male fitness and mating in Qfly males and females. These resources will further help facilitate future research on genetics of mating and fitness, which will ultimately provide longer term benefits for Qfly SIT.

The postdoctoral fellow led a further fruit fly study while employed on the HIA grant – a desk-based study collating information on the niche over-lap of *B. tryoni* and its closely related sibling species *B. neohumeralis*. While at first seemingly removed from the primary grant, the collation of the phenotypic data for the two species is essential when trying to interpret the nearly 100% over-lap in the genomes of the two species as documented by Gilchrist et al. (2014) – published at the very start of the fellowship. This study reinforced the need to continually link phenotypic studies with genotypic studies.

Methodology

Objective 1. Training of a post-doctoral scientist

Dr Kumaran Nagalingam was the postdoctoral scientist mentored and supported by this project. Kumaran completed his PhD in the QUT Fruit Fly group, and undertaking the transformational postdoctoral fellowship in the same lab had the potential to be 'just more of the same', rather than a true postdoctoral experience. To overcome this potential problem, which was identified from the very start of the grant, his primary PhD supervisor, Prof Tony Clarke, played only an 'overseeing' role in the project, to ensure the research remained aligned with HIA approved project objectives and to provide a high level, personal mentoring role to Kumaran. But for nearly all day-today activities, QUT academic staff with whom Kumaran had only minor or no contact during his PhD, notably specialist geneticists Dr Peter Prentis, Dr Kevin Dudley and A/Prof Stephen Cameron, played active research and mentoring roles for Kumaran, giving him technical training, genetics education and professional mentoring, which was quite different to his PhD experience. Note that only Cameron was named on the original grant, yet Prentis and Dudley have also provided very significant FTE to the project, not only in mentoring and training, but also in direct research and analysis time. Stephen Cameron left QUT in mid-2016 to take up a new position in the U.S.A., but has remained involved in the project via electronic communication. Discussion with HIA at the time of Stephen's departure determined that there was no need to make a formal project variation with respect to staffing.

Objective 2. Developing a deeper understanding on the genetics of male fitness

2a. Heritability of lure foraging trait

Background: The effects of plant secondary compounds on *Bactrocera* fruit flies are complex. The plant derived compounds (= male lures) change the genetic expression and energy metabolism of flies that ultimately render flies fitter (Kumaran et al. 2014). In addition, in Qfly, offspring sired by males fed on lures show higher lure foraging ability than the offspring of non-lure fed fathers (Kumaran & Clarke 2014). The lure effect is indirect in the offspring from the father through the mother, i.e. the effect is inherited in the offspring. In this project, which we built form our earlier work, we investigated more closely the heritability of the lure foraging trait using Qfly culture lines from Brisbane and Camden. We used classical genetic selection trails, selecting the most rapidly lure responsive males from lure-fed lines across nine generations. Subsequent to the selection trial, we performed RAD-sequencing to identify single nucleotide polymorphism (SNPs) in the F9 lure selected lines compared to the parental (i.e. starting) populations.

Methodology: The classical selection trials were carried out at the Samford Ecological Research Facility in large field enclosures. For this, 14-day old males were fed with cue lure and crossed with virgin mature females to obtain offspring sired by lure-fed males. Females were provided with oviposition substrates and eggs were collected and reared through to adults. The emerged adults, which served as F1 offspring of lure-fed males were used for testing the foraging ability. Two-hundred-and-forty F1 offspring males from lure-fed lines were released into large field cages when they are sexually mature (12-14 days old) and presented with cue lure. Observations were made on the number and timing of flies arriving at the lure sources following the protocol in Kumaran and Clarke (2014). For the treatment (i.e. lure exposed) lines, the first 25% of flies responding to and feeding on the lure source were collected to be used as the sires for the next generation, likewise for nine generations. Thus, a positive

selection for lure response at each generation through to the F9 was achieved. A separate control line of offspring sired by non-lure fed males were also obtained for comparison by using randomly selected non-lure exposed flies as sires for the next generations. The control lines were run for their lure response for each of nine generations at the same time as the treatment lines.

To develop supportive genetic data on foraging ability before and after selection, tissues from the parental lines and F9 offspring lines were subjected to Restriction site associated DNA sequencing (RAD-seq). The objective was to develop genetic data that helped identify the genetic basis of the heritable lure effects through identification of SNPs in the lure selected lines. Single Nucleotide Polymorphism (SNP) is variation in a single nucleotide that occurs at a specific position in the genome (Morin et al. 2004), and RAD-seq approach is known to effectively estimate the heritability of complex traits associated with fitness of an organism, similar to lure feeding in tephritids, through identifying SNPs (Baird et al. 2008).

2b. Epigenetic modifications after lure feeding in Qfly

Background: The objective of this study was to determine if modifications in histone proteins occurred after male Qfly fed on cue lure. The technique used was the ChIP-seq (Chromatin immunoprecipitation-sequencing) approach, which dissects out possible epigenetic effects of lure feeding by mapping the histone modifications in lure-fed males against control males. Histones are the chief protein components of chromatin, which mediate DNA functions, and play a significant role in the regulation of gene expression (Jaenisch & Bird 2003). We believed the approach could potentially unveil the factors that facilitated the differential expression of functional genes after lure feeding in Qfly (as identified in Kumaran et al. 2014).

Methodology: ChIP includes crosslinking the bound protein (DNA) with chromatin, lysing the chromatin extractions and fragmentation of the chromatin (by sonication to generate chromatin fragments of 200-1000 bp size) (Thorne et al. 2004). Then the chromatin is immuno-precipitated through conjugation using specific antibodies that can recognize the bound proteins or protein modifications of interest (Park 2009). Finally, crosslinks are reversed and the free DNA is sequenced to determine the sequences bound by the protein (Nelson et al. 2006). For Qfly ChIP-seq, flies were fed on cue lure and their head tissues were crosslinked using formaldehyde to link the chromatin with DNA. For immuno-precipitation, seven functional antibodies *viz*., H3K36me3, H3K36me1, H3K4me3, H3K4me2, H3K27me3, H3K27acetylation and Histone 3 (nuclear loading control) were conjugated and immuno-precipitated with chromatin using ChIP grade reagents. After conjugation, chromatin was eluted and quality checked prior to library preparation. The libraries were sequenced using Illumina Next-seq 500 and high quality sequences were obtained for three *B. tryoni* lines (Boggo Road, Camden and Cairns populations). A detailed methodology is explained in the appended draft manuscript as this process was almost entirely novel, and was certainly the first time done for any tephritid fruit fly (Appendix 1, ChIP methodology manuscript).

2c. Comparative transcriptome analysis of females mated with lure-fed males unfed males

Background: Transcriptome profiles are the functional resources to understand the genetic mechanisms mediating the phenotypic changes in any organism. For instance, stage specific transcriptome profiling has been done in several insect species to identify the transcription factors regulating sexual development and reproduction (Gomulski et al. 2012; Zheng et al. 2012). Only very few studies on female specific transcriptional regulation have been done previously in tephritid fruit flies and other systems (McGraw et al. 2008; Gonclaves et al. 2013); in those studies while relatively very

few genes were found differentially regulated in females post mating, the results nevertheless provided significant insights on female mating in *Anastrepha obliqua* and *Drosophila melanogaster*.

In the current study, using RNA-seq analyses, we compared virgin and mated Qfly females to understand the transcriptional factors regulating mating, but notably, we also compared females mated with normal males and lure-fed males to profile the transcriptional factors regulating female re-mating. We have previously shown that females mated with lure-fed males re-mate less frequently than females mated with lure unfed males (Kumaran et al 2013). Thus difference in the gene transcripts of the two categories of females may be related to female propensity to re-mate.

Methodology: Three groups of females were maintained: mature virgin females, females mated with lure unfed males, and females mated with males fed on the male lure zingerone. A total of 20 females (14 days old) from each of the three female groups were collected and snap-frozen in liquid nitrogen the morning after the flies had mated with the respective males. Total RNA was extracted from whole bodies using Trizol and purified with a Qiagen RNeasy kit following the manufacturer's instructions. Detailed methodology on isolation, library construction, assembly and annotation are presented in Arthofer et al. (2014).

To determine genes that were differentially expressed (DEGs) between virgin females and mated females (mated with normal males), and between females mated with lure-fed and unfed males, P value threshold for the analysis was determined using false discovery rate (FDR). We undertook gene set enrichment analysis to determine whether particular GO categories were overrepresented in the DEGs. The detailed methodology is explained in the appended draft manuscript (Appendix II) and Kumaran et al. (2014).

Outputs

Objective 1. Trained a postdoctoral scientist

The genetics research was undertaken within the Genomics Laboratory of QUT's Central Analytical Research Facility (https://www.qut.edu.au/research/research-projects/genomics-laboratory), a QUT centrally run and supported facility which is equal or better to any genomics laboratory in Australia. With the guidance of his mentors, Kumaran Nagalingam was trained to use advanced techniques, such as chromatin immunoprecipitation and restriction site associated DNA polymorphism sequencing, to understand genetic and epigenetic modifications in Qfly, in addition to training on the interpretation and analysis of classical genetic selection trials.

The specifically identified conference travel money that was provided within the grant was used to support Kumaran to undertake a professional development trip of approximately four weeks, which took in two conference/workshop presentations and two sets of professional meetings. From late June to mid-August 2016, Kumaran presented research findings at the International Society for Behavioural Ecology Conference in Exeter (27th July – 3rd Aug) and at the first Tephritid Workers of Asia, Australia and Oceania conference in Malaysia (15th-18th Aug). Between the two meetings he met and spoke with colleagues at Oxford University, and then spent a full week at the IAEA Laboratories in Seibersdorf, outside of Vienna, meeting with colleagues who work full time on the genetics of SIT flies.

The only obvious weakness of the fellowship was that no papers were published during its course. This was simply due to the technical complexity of the work undertaken, the immense amount of new data generated, and the time it takes to analyze and consider that much data. That all experiments were completed, all data generated, and nearly all data fully analyzed is an excellent outcome from a two-year position. That three out of four manuscripts anticipated from this postdoc are presented in draught form with this final report demonstrates that the submission of these manuscripts will occur within the next few months.

While difficult to judge the success of a postdoctoral experience, the senior members of the project team are confident that Kumaran received as strong a postdoctoral experience as was possible to provide within the two years. That Kumaran was offered a prestigious CSIRO postdoctoral position prior to the completion of the current fellowship shows that his expertise as an emerging applied entomologist is externally recognized.

Objective 2. Study the heritability of the lure foraging trait and genetic mechanisms

Our primary aim was initially to generate data for two high quality publications. However, we have subsequently studied heritability of lure foraging, epigenetic modifications of lure feeding, and differential gene expression in females mated with lure-fed and unfed males and thus, in two years, generated data for three independent manuscripts that are under preparation to submit to journals (drafts of two manuscripts appended). In addition, a manuscript on the comparative ecology of Qfly and lesser Qfly was developed using desktop analysis of historical and contemporary data sets (Appendix III).

Manuscript 1. Chromatin immunoprecipitation in tephritid fruit flies - methodology

Interactions between DNA and underlying proteins drive the transcriptional regulation and expression of

genes (Felsenfeld & Groudine 2003). Modifications in histone proteins (referred to as the epigenome) play a key role in functional gene regulation and expression and hence regulate crucial biological processes (Peterson & Laniel 2004; Kouzarides 2007). The technique predominantly used for epigenome profiling is chromatin immuno-precipitation (ChIP) followed by identification of DNA fragments through sequencing (ChIP-seq): critical steps for effective ChIP are (i) cross-linking (ii) shearing to the right size (iii) immuno-precipitation and (iv) DNA isolation.

The ChIP technique has been used in several model species to understand complex biological processes (Furey 2012). However, to the best of our knowledge, there are no ChIP studies on non-model species, and this is perhaps because the methodology developed for model species often fail to work for non-model species, or at least requires considerable alterations. We tested ChIP methodology described by manufacturers (Abcam) and a methodology followed for the model species *Drosophila* (Tran et al. 2012), in our non-model tephritid fruit flies. These methods failed to effectively immuno-precipitate chromatins and resulted in poor recovery of quality chromatin DNA for downstream analyses because of ineffective cross-linking and poor antibody conjugation.

After multiple attempts of protocol optimization for more than a year, we successfully developed a working protocol to study epigenetic modifications in non-model fruit flies. The protocol was developed combining the manufacturer's guidelines and previous studies on model organisms. We checked the chromatin extractions from this modified protocol for quality and found that chromatin is cross-linked effectively and shearing (sonication) yielded the right size product for downstream applications. We checked the conjugated antibodies and immuno-precipitation by Qubit assays and the result of the conjugation was confirmed with Qubit and bioanalyzer profiles. Overall, the methodology delivered successful sequencing of the chromatin DNA with high quality reads for downstream analyses. The Diffbind analysis of histone proteins also evidenced histone modifications in Qfly and hence confirmed that the methodology has worked effectively for Qfly. This is a first report on histone modifications in tephritid fruit flies, and the manuscript explaining the detailed methodology is being prepared to submit to the journal 'Genomics' (draft appended, Appendix I).

Manuscript 2. Heritability of lure foraging trait and epigenetic modifications after lure feeding

Heritability of lure foraging trait was studied using classical half-sib selection trials and state-of-the-art genotyping by sequencing technique. Classical selection trails in large field cages for nine generations demonstrated that the flies selected for the greater lure foraging ability sired offspring with yet greater lure foraging ability, and showed that we could potentially select for male lines that are more effective in foraging for and finding lure sources. Figures 1 & 2 demonstrate the increase in the number of flies responding to a cue lure source over nine generations when compared with the control lines of same generation and the positive selection of the lure foraging trait in the lure selected lines. Note that generations 4 to 6 are anomalous with the rest of the data set because field cage trials needed to be run during winter and there were generally very slow behavioral responses from both selected and control lines.

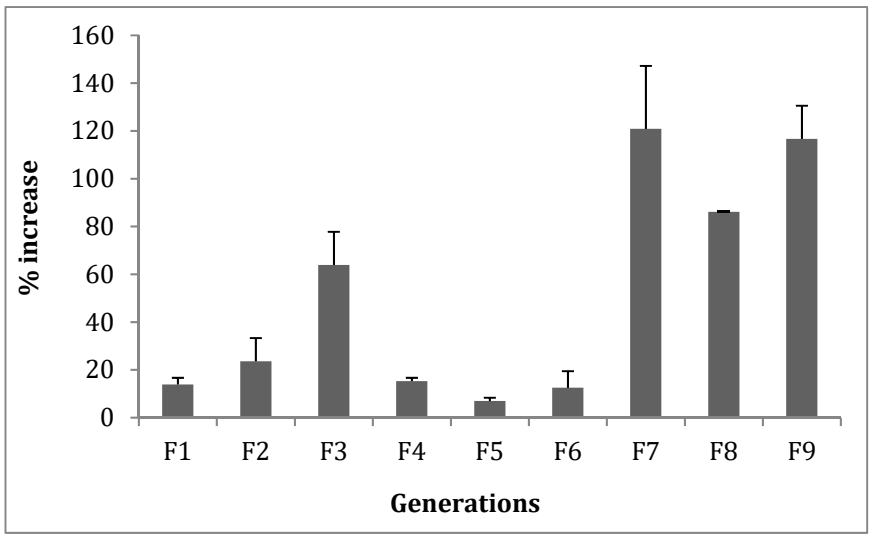


Figure 1. Percent increase in male *Bactrocera tryoni* response to cue lure source over generations during the first 15 minutes of lure exposure after positive selection for rapid lure foraging across generations.

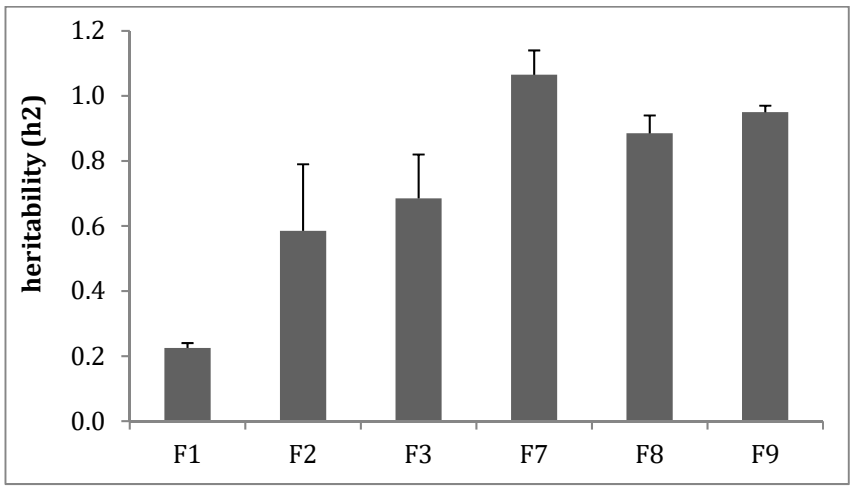


Figure 2. Heritability (h2) calculated over generations using total response of parental populations, response by parental lines (top 25% males) and offspring response.

Genotyping-by-sequencing was done using the parental lines, F9 selected lines and F9 control lines. We generated very large datasets from this project that will be used to identify SNPs in the selected lines compared with parental and control lines. The sequencing of GBS analysis yielded 49.12 Gb quality data with 327,493,943 quality reads for potential SNPs identifications (Table 1). This will add value to the genetic resources already available to assist Qfly management.

Table 1. A snapshot of GBS sequencing output from parental, lure-selected and control lines.

Lane	Sample Name	Single Reads	Data Yield (bp)
150bp Single End - Flowcell ID: HHC5JBGXY			
1	7096-1A	22,052,171	3.31 Gb
	7096-1B	18,637,774	2.80 Gb
	7096-2A	21,230,964	3.18 Gb
	7096-2B	20,641,593	3.10 Gb
2	7096-1A	21,308,803	3.20 Gb
	7096-1B	18,090,467	2.71 Gb
	7096-2A	20,582,233	3.09 Gb
	7096-2B	20,057,589	3.01 Gb
3	7096-1A	22,374,592	3.36 Gb
	7096-1B	18,868,446	2.83 Gb
	7096-2A	21,496,383	3.22 Gb
	7096-2B	20,918,228	3.14 Gb
4	7096-1A	21,695,834	3.25 Gb
	7096-1B	18,343,878	2.75 Gb
	7096-2A	20,871,607	3.13 Gb
	7096-2B	20,323,381	3.05 Gb
Total		**327,493,943**	**49.12 Gb**

The histone modification study was conducted using the ChIP-seq approach. The results of the ChIP-seq analysis confirmed histone modifications in Qfly for the first time. Overall, we have identified more than 20000 peaks that confirm possible histone modifications in Qfly. In flies fed with lures we identified ~200 peaks for H3K4me3 and ~160 peaks for H3K36me3 antibodies differentially expressed in lure-fed males when compared with normal males through Diffbind analyses. Further analysis with other antibodies *viz*., H3K36me1, H3K4me2, H3K27me3 and H3K27acetylation is being carried out. The enormous amount of genetic data developed from sequencing of Camden, Brisbane and Cairns lines will be an added resource to study underlying epigenetic mechanisms regulating mating and reproduction in Qfly and other tephritids.

A draft manuscript of this work is not presented in this report because the heritability analysis of individual SNPs from the GBS analysis is still underway. The analysis has been out-sourced to a commercial provider to ensure rapid completion (and subsequent publication), now that the postdoctoral fellow has moved to a new position.

Manuscript 3. Intrinsic female transcription factors regulating female mating and re-mating

Regulation of female re-mating by male genotype through indirect genetic effects (IGEs) is well understood as there is an extensive empirical work on post-mating physiological modulation in females by male accessory gland proteins and sperm (Gillott 2003). Surprisingly, the role of female specific factors in regulating mating and re-mating remain largely unresolved, even though females are the decisive gender in choosing whether to mate or not, and with whom (Andersson & Iwasa 1996). Hence, in this study, we compared virgin and mated females to understand female transcriptional factors regulating mating, but most importantly, we compared females mated with normal and lure-fed males to understand transcripts regulating re-mating and other post mating changes using comparative RNA-seq analyses.

The results revealed gene transcripts possibly regulating mating and re-mating in females. Figure 3 is a heat-map of transcriptome profile showing differential expression of genes with a group of genes up-regulated (yellow pattern) and down-regulated (purple pattern) in females when mated with lure-fed males. There were 89 genes differentially expressed (DEGs) in females when mated with lure-fed males, with 70 genes up-regulated and 19 genes down-regulated. Investigation of gene categories revealed enrichment of several GO terms. There were 126 GO categories enriched within the biological processes GO term, six in the cellular processes GO term, and 37 in the molecular function GO term. The detailed results on the number of genes up- and down-regulated are incorporated in the draft manuscript (Appendix II). The down-regulated gene transcripts in females mated with lure-fed males are perhaps the regulating factors of female re-mating, since the females mated with lure-fed males showed reduced re-mating propensity in our earlier studies. We strongly believe that the genetic information generated through transcriptome analyses help understand the factors driving females to mate and re-mate.

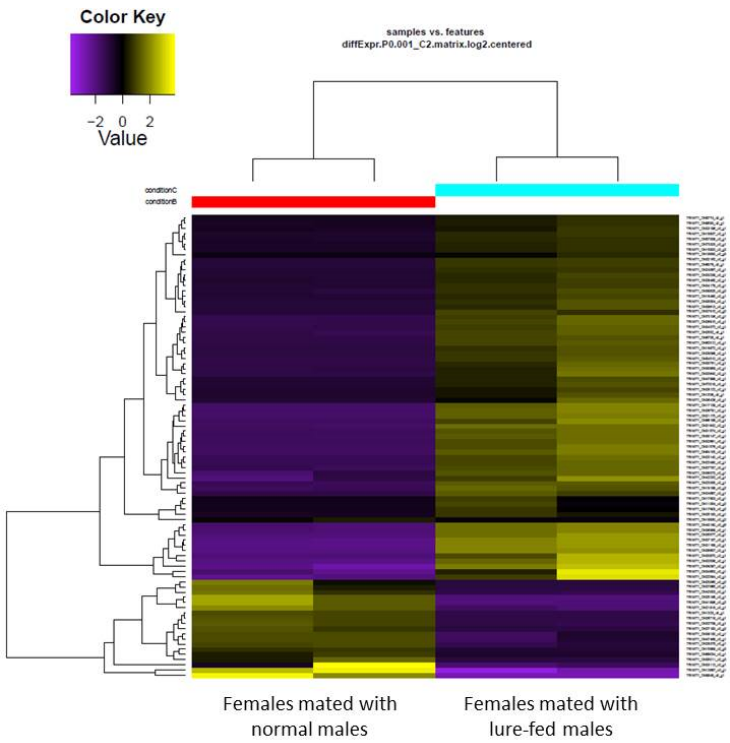


Figure 3. A heat-map of Qfly female transcriptome profiles after mating with lure-fed and lure-unfed males

Manuscript 4. No evidence of niche segregation in *Bactrocera neohumeralis* (Diptera: Tephritidae), a sibling to the Queensland fruit fly, *Bactrocera tryoni*

The Australian tephritid species *B. tryoni* and *B. neohumeralis* present a unique example for possible sympatric speciation. These two sibling species exhibit very little variation in their morphology: *B. neohumeralis* has a brown numeral calli, which is yellow in *B. tryoni*. Genetically, microsatellite analyses reveal no differences between these two species or the difference is trivial (Wang et al. 2003; Gilchrist & Ling 2006), while their entire genomic overlap is much greater than that seen between populations of some species (Gilchrist et al. 2014). In terms of their mating behaviour, *B. neohumeralis* copulate in bright light during the middle of the day, while Qfly mates at dusk when light intensity is low (Wolda 1967; Pike & Meats 2002): copulation time is the only functional variation known to be maintaining reproductive isolation between these two species. While there is a large amount of comparative genetic and mating data on these two species (Clarke et al. 2011), only one study (Gibbs 1967) has previously directly compared the two species in the field, and this was done over only one season at one site (Rockhampton). Understanding differences and similarities in *B. tryoni* and *B. neohumeralis* is important, as despite their extremely close genetic similarity, Qfly is a major pest, while *B. neohumeralis* is locally significant pest only, with its pest status restricted largely to central and northern Queensland. To better understand these species, we compared abundance of both species in different landscapes, their seasonal abundance over several years, and abundance in different host fruits to investigate for the evidence of niche separation, if any, between *B. neohumeralis* and *B. tryoni*.

Abundance of *B. neohumeralis* and *B. tryoni* was highly correlated and followed very similar phenologies over time. Both species followed a similar seasonal variation in all locations studied, with greater numbers of flies trapped during Sep-Nov and Feb-March in most of locations. With respect to habitat use, there were greater number of flies trapped in dry sclerophyll forests followed by horticulture farming systems, and this habitat use pattern was also similar in both *B. tryoni* and *B. neohumeralis*. Host use did not differ between *B. tryoni* and *B. neohumeralis*, with similar patterns of increase and decrease in fly populations infesting multiple hosts. Overall, the data showed no evidence of niche segregation in *B. neohumeralis* and *B. tryoni*, and fail to explain the increased pest status of Qfly over *B. neohumeralis*. The draft manuscript of this paper is presented as Appendix III.

Recommendations for strategic research which could be applied to improving the quality of sterile flies

This project has developed a very large amount of *B. tryoni* genetic data related to male fitness, female mating and female re-mating (Table 2). The data on lure foraging, and genetic and epigenetic mechanisms mediating male fitness provide preliminary information on genetic and epigenetic changes occurring in the Qfly males after lure feeding. Further systematic studies targeting the specific genes and gene families modified/regulated after lure feeding are warranted to develop potentially fitter male lines for SIT. In addition, the data on female transcriptome provided genes and pathways modified when mating with lure-fed males. Further targeted studies on the gene transcripts that are down-regulated when females mate with lure-fed males could tell us the genes the need to be knocked down to avoid female re-mating. For instance, if we can develop fitter male sterile lines that can knockdown the genes that trigger female re-mating, then such fitter sterile lines while competing more effectively with wild males could render females unreceptive for further matings: that ultimately increases the odds of the SIT success.

Table 2. Genetic data generated from all three studies

	yield	%>Q30	Reads
ChIP-seq (epigenetics)			
Run 1 (Batch 1 Test)	10.67Gb	87.83	133,701,100
Run 2 (Batch 2 Test)	18.99Gb	95.86	244,033,044
Run 3 (Batch 1)	39.85Gb	95.22	516,820,832
Run 4 (Batch 1 Set 2)	53.96Gb	92.01	747,981,024
Run 5 (Batch 3)	36.42Gb	95.83	463,868,696
Run 6 (Batch 2)	35.78Gb	95.3	457,484,192

	yield	%>Q30	Reads
Female transcriptome	38.72Gb	95.21	546,057,680

	yield	Reads
GBS (RAD-seq)		
Lane 1	12.39Gb	82,562,502
Lane 2	12.01Gb	80,039,092
Lane 3	12.55Gb	83,657,649
Lane 4	12.18Gb	81,234,700
Total	**49.13Gb**	**327,493,943**

Outcomes

Objective 1. A highly qualified research scientist with capacity to significantly aid the national effort to make the SIT a viable and sustainable control method for Qfly

A highly qualified research scientists with expertise in fruit fly epigenetics, transcriptomics and genomics, with international and domestic connections to fruit fly researchers, has received advanced training and mentoring. The research undertaken will help make the SIT a sustainable control method for Qfly. An ongoing position in fruit fly genetics was not available and the fellow is now working in applied entomology within the CSIRO Weeds Biological Control program where his skills are being used to help manage weeds of national importance.

Objective 2. New research knowledge which will improve the SIT, making the technique more viable as a sustainable control strategy

Heritability of lure foraging trait

This project has found that the lure foraging trait in Qfly appears heritable, and provided a first line of evidence that lure feeding by males is modifying the underlying genotype of the offspring indirectly through females that are mating with the lure-fed males. The offspring sired by the males (father) with a greater lure foraging trait were found to inherit the foraging ability of their fathers. In addition, this project suggests that we can develop culture lines with a greater ability to forage for natural lure resources.

The genetic data developed in addition to the classical selection trial data will identify SNPs in the male lines selected for greater lure foraging ability. Identification of SNPs will be our next target for the deeper understanding of fitness related genetic components mediating physical fitness and mating in male Qfly. This will help overcome problems with the SIT associated with poor quality, competitively weaker male flies.

Genetic mechanisms mediating male fitness

We have generated supporting genetic knowledge (both genomic and epigenomic) that can help future studies to thoroughly understand male fitness in Qfly. The epigenetic data, in particular, provides essential knowledge on in-depth mechanisms mediating male fitness in Qfly, particularly by identifying genes and genome regions related to physical fitness of males and mating success. For the first time in tephritid fruit flies, we have generated data associated with the epigenome of the flies, with possible overall histone modifications and modifications specific to reproductive fitness. The histone modifications found provide insight (and areas for targeted research) on the genotypic factors regulating the various biological functions and processes in Qfly as well as in other tephritid fruit flies.

Genetic mechanisms mediating female mating and re-mating

This part of the project found genes differentially expressed in mated females compared with unmated females. There were several genes upregulated in mated females, suggesting those genes are perhaps regulating the mating processes and post-mating physiology in female *B. tryoni*. Further, we compared females mated with normal males and lure-fed males. This comparison was done to investigate genes regulating the post-mating physiological changes observed in our previous studies in Qfly females. The results revealed several genes and pathways differentially expressed in females mated with lure-fed males. Future studies should target the genes up- and down-regulated within the DGEs to better

understand the genomic basis of female post-mating physiology. For instance, the down regulation within the DGEs of females mated with lure-fed males in particular is important because those might be the genes mediating female re-mating. Should further studies confirm the role of DGEs in regulating female re-mating through knockdown, then it is probable that we can search for a genetic mechanism in males that can suppress those genes in females through induced indirect genetic effects. This will ultimately help avoid a larger proportion of females re-mating with wild fertile males after being first mated with sterile males.

Summary

Overall, the project has delivered genetic knowledge on factors that are mediating mating success in Qfly. While the results show changes in the genetic makeup of flies (in flies fed with lures, females mated with lure-fed males and offspring sired by lure-fed males), the information generated can be essentially applicable to general male fitness and female mating patterns regardless of lure feeding. The epigenetic modifications observed in the lure-fed males provide genotype knowledge which needs to be targeted to improve male fitness. Likewise, the female transcriptome shows us genes that may be regulating female mating and re-mating. The heritability of lure foraging trait and genotyping-by-sequencing results provide significant genetic information for future studies that identify novel candidate gene(s) for active manipulation through quantitative trait loci (QTL) mapping that will help us induce the lure effect without actually exposing flies to lures. Together, this knowledge provides a genetic basis for developing male SIT lines which can mate successfully with wild females and can probably switch off the genes and pathways that induce re-mating in females, ultimately resulting in more effective SIT control of Qfly.

Evaluation and Discussion

Project delivery

As a transformational postdoctoral fellowship, this project developed a high quality scientist with advanced skills and experience in cutting-edge technologies. While developing genetic resources to understand the complex behaviors in Qfly, this project, we believe, has transformed a doctorate student into independent scientist with knowledge on fruit fly behavioral genetics that ultimately facilitate the Qfly pest management through SIT.

The project was effective in achieving its aims with significant outputs and outcomes. The project generated a vast amount of genetic data (Tables 1 & 2) that will be an invaluable addition to the resources that help understand fitness components of the Qfly. In addition, we developed a methodology to study histone modifications in tephritid fruit flies to assist future studies. Specifically, the studies on epigenetic modifications have identified broad-spectrum histone modifications in Qfly, and explicit modifications due to lure feeding. This outcome will greatly help future projects to target the genetic mechanisms that are possibly mediating physical and mating fitness in *B. tryoni*.

While this project has generated more research knowledge than aimed, we cannot conclude that the project has been completed in all aspects because the papers have not yet been submitted. Two years was a very short time for project and we have developed genetic approaches and data that normally might be expected to take 4 or 5 full years. The down-side of this was that the time between data collection/analysis and project completion was not enough to allow full manuscript preparation. We believe the project is successful in terms of research knowledge generated and the likely impact the data will create among research communities. We are aiming to have three high quality publications using the data generated, among which we are aiming to publish one of the publications in the very top tier journals such as *Nature*, *Science* or *Molecular Biology and Evolution*. We are justified in trying for such journals, as we have, for the first time in any system, identified the genetic mechanisms for run-away sexual selection which is of great theoretical importance in biology.

This transformational postdoctoral fellowship was aimed at understanding the genetic basis of fitness traits in Qfly, for the subsequent development of high quality SIT lines with greater physical and mating fitness. The knowledge we have gained will be presented in the scientific literature for other researchers (the next-users of this data). The data sets themselves will be made fully and publically available at the time of paper acceptance, as is normal practice in this field.

Project-team self-evaluation

The aim of the transformational postdoctoral fellowship was to train and develop a high quality scientist while generating the genetic and behavioural data related to Qfly fitness.

We believe we have done this. We introduced cutting-edge concepts such as epigenetics and genotyping-by-sequencing to fruit fly research, so helping to understand the genetic mechanisms mediating various biological processes. We simultaneously trained a scientist to use these cutting-edge skills. We have generated high quality data pertaining to genetic basis of Qfly fitness that will help future targeted research work that are now working to develop quality sterile lines within the SIT Plus consortium. The data and results were discussed with peers through conferences and especially with peers at the International Atomic Energy Agency and received positive feedback.

Given the limited time available (2 years) to complete what became a very large genomics project, the manuscripts generated are still very early drafts. While we are entirely happy with the data generated and training provided to the postdoctoral fellow, we are not equally happy with the publication side of the project, which we will continue to pursue and finalise over the next four to six months.

Learning from the project

The project was absolute discovery research with much significant work needed to advance the research area before seeing applied benefits. In addition to developing a research scientist, part of the project (the actual research part) focused on developing more genetic resources to assist the future fruit fly research work targeting SIT. We strongly believe the research area needs to advance with future funding either through HIA or other relevant funding sources.

Overall relevance to industry

Given that there is national drive to implement SIT for *B. tryoni* control in Australia, the project is of high relevance to the horticultural sector as it provides specific genomic targets to improve the quality of sterile male breeding lines within the SIT factory. The work was a discovery/training program, however, and direct benefit to growers is still several years away, and entirely dependent on further research and development.

Recommendations

This project aimed at developing a qualified research scientist with capacity to significantly assist the national effort to make the SIT a viable and sustainable control method for Qfly; and to generate research knowledge to help future projects identify gene targets to make the SIT flies fitter.

Our recommendations to progress these areas are:

1. Consideration of long-term funding for specialist researchers who, having been supported to develop a unique skill set directly relevant to a national horticultural priority, now need to move to other sectors (in this case weeds) to remain employed.
2. Consideration be given to classical selection as a relatively easy method of reinforcing preferred traits into factory brood lines.
3. Genomics researchers target the individual genes and gene ontology groups identified in this project to make more rapid advancement in the genomic selection of fitter male lines.

Scientific Refereed Publications

Journal articles

Kumaran, N., Dudley, K. Lorenc, M., Manoli, S. Cameron, S. L. and Clarke, A. R. Chromatin immuno-precipitation and histone modifications in non-model fruit flies (Diptera: Tephritidae). Genomics. In Prep (preliminary draft appended Appendix I).

Kumaran, N., Dudley, K. Prentis, P. J., Cameron, S. L. and Clarke, A. R. Runaway selection and the sexy-son hypothesis explained by genetic and epigenetic modifications. Nature/Molecular Biology and Evolution. In Prep. (Draft not appended as the results need further statistical and bioinformatics analyses)

Kumaran, N. Prentis, P. J., Chloe, W. D. B., Xin-Ji. and Clarke, A. R. Intrinsic female transcription factors regulating mating and post mating physiology. Genome Biology and Evolution. In prep (preliminary draft appended Appendix II).

Kumaran, N., Veuti, E., Missenden, B. Senior, L., Higgie, M. and Clarke, A. R. No evidence of niche segregation in *Bactrocera neohumeralis* (Diptera: Tephritidae), a sibling to the Queensland fruit fly, *Bactrocera tryoni*. Bulletin of Entomological Research. In prep (preliminary draft appended Appendix III).

Table 3. Scientific presentations and discussions

Speaker	Title of the presentations/ meetings	Meeting and date
Anthony Clarke	The 'Red Bull' effect: an additional explanation for response to phenylpropanoids by male *Bactrocera* fruit flies.	International Society of Chemical Ecology (ISCE) Conference 2015, 29th June to 3rd July, Stockholm, Sweden
Anthony Clarke	Orchids and fruit fly pest management: what's the link?	Departmental seminar, University of Würzburg *Biozentrum*, 21st April 2016, Würzburg, Germany
Kumaran Nagalingam	Sexual selection in *Bactrocera* fruit flies	International Society for Behavioural Ecology (ISBE) Conference 2016, 27th July to 3rd August, Exeter, UK
Kumaran Nagalingam	Meeting and discussions with Oxford University staff	Oxford University, 2016, 4th and 5th August, Oxford, UK
Kumaran Nagalingam	Heritability of lure foraging trait and functional role of male lures	International Atomic Energy Agency, 2016, 8th August to 12th August, Vienna, Austria
Kumaran	Effect of Male Lures on Pheromone Release	Tephritid workers of Asia, Australia and

Nagalingam	and Energy Metabolism in Queensland Fruit Fly, *Bactrocera tryoni*	Oceania conference (TAAO), 2016, 15th to 18th August, Putrajaya, Malaysia
Kumaran Nagalingam	Functional role of male lures of *Bactrocera* fruit flies: potential to maximize their use in pest management (Invited talk)	Entomological Society of Queensland, 13th September 2016, Ecosciences Precinct, Brisbane, Queensland, Australia

Intellectual Property/Commercialisation

No commercial IP was generated through this project. New knowledge developed was shared to the SIT Plus consortium through direct briefings and sending of progress reports. The data are currently being prepared to publish in the scientific journals and once done the genetic material will be placed on publically accessible data bases.

References

Andersson M. & Iwasa Y. (1996). Sexual selection. *Trends in Ecology & Evolution* 11, 53-58.

Arthofer W., Clarke A.R., Kumaran N., Prentis P.J., Schafellner C., Schlick-Steiner B.C. & Wachter G. A. (2014). Genomic resources notes accepted 1 December 2013–31 January 2014. *Molecular Ecology Resources* 14, 664-665.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A. & Johnson E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* 3, e3376.

Calkins C.O. & Parker A.G. (2005). Sterile insect quality. Pp 269-296 in *Sterile Insect Technique*. *Principles and Practice in Area-Wide Integrated Pest Management* (V.A. Dyck, J. Hendrichs & A.S. Robinson eds). Springer, Dordrecht, The Netherlands.

Clarke A.R., Powell K.S., Weldon C.W. & Taylor P.W. 2011. The ecology of *Bactrocera tryoni* (Diptera: Tephritidae): what do we know to assist pest management? *Annals of Applied Biology* 158, 26-54.

Dyck, V.A., Hendrichs, J., & Robinson, A.S. (eds) (2005). *Sterile Insect Technique*. *Principles and Practice in Area-Wide Integrated Pest Management* . Springer, Dordrecht, The Netherlands.

Felsenfeld G. & Groudine M. (2003). Controlling the double helix. *Nature* 421, 448-453

Furey T.S. (2012). ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nature Reviews Genetics* 13, 840-852.

Gibbs G.W. (1967). The comparative ecology of two closely related, sympatric species of *Dacus* (Diptera) in Queensland. *Australian Journal of Zoology* 15, 1123–1139.

Gilchrist A.S. & Ling A.E. (2006). DNA microsatellite analysis of naturally occurring colour intermediates between *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy) (Diptera: Tephritidae). *Australian Journal of Entomology* 45, 157-162.

Gilchrist A.S., Shearman D.C.A., Frommer M., Raphael K.A., Deshpande N.P., Wilkins M.R., Sherwin W.B. & Sved J.A. (2014). The draft genome of the pest tephritid fruit fly *Bactrocera tryoni*: resources for the genomic analysis of hybridising species. *BMC Genomics* 15, 1153.

Gillott, C. (2003). Male accessory gland secretions: modulators of female reproductive physiology and behavior. *Annual Review of Entomology* 48, 163-184.

Gomulski L.M., Dimopoulos G., Xi Z., Scolari F., Gabrieli P., Siciliano P., Clarke A.R., Malacrida A.R. & Gasperi G. (2012). Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitis capitata*. *PloS One* 7, e30857.

Gonçalves V.R., Sobrinho I.S., Malagó W., Henrique-Silva F. & Brito R.A. (2013). Transcriptome analysis of female reproductive tissues of *Anastrepha obliqua* and molecular evolution of eggshell proteins in the fraterculus group. *Insect Molecular Biology* 22, 551-561.

Henning G.T., Schild S.E., Stafford S.L., Donohue J.H., Burch P.A., Haddock M.G., Trastek V.F. & Gunderson L.L. (2000). Results of irradiation or chemoirradiation following resection of gastric adenocarcinoma. *International Journal of Radiation Oncology*Biology*Physics* 46, 589-598.

Jaenisch R. & Bird A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics* 33, 245-254.

Kouzarides T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.

Kumaran N., & Clarke A.R. (2014). Indirect effects of phytochemicals on offspring performance of Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae). *Journal of Applied Entomology* 138, 361-367.

Kumaran N., Balagawi S., Schutze M.K. & Clarke, A.R. (2013). Evolution of lure response in tephritid fruit flies: phytochemicals as drivers of sexual selection. *Animal Behaviour* 85, 781-789.

Kumaran N., Prentis P.J., Mangalam K.P., Schutze M.K. & Clarke A.R. (2014). Sexual selection in true fruit flies (Diptera: Tephritidae): transcriptome and experimental evidences for phytochemicals increasing male competitive ability. *Molecular Ecology* 23, 4645-4657.

Mankin R. W., Lemon M., Harmer A.M., Evans C.S. & Taylor P.W. (2008). Time-pattern and frequency analyses of sounds produced by irradiated and untreated male *Bactrocera tryoni* (Diptera: Tephritidae) during mating behavior. *Annals of the Entomological Society of America* 101, 664-674.

McGraw L.A., Clark A.G. & Wolfner M.F. (2008). Post-mating gene expression profiles of female *Drosophila melanogaster* in response to time and to four male accessory gland proteins. *Genetics* 179, 1395-1408.

Meats A., Holmes H.M. & Kelly G.L. (2004). Laboratory adaptation of *Bactrocera tryoni* (Diptera: Tephritidae) decreases mating age and increases protein consumption and number of eggs produced per milligram of protein. *Bulletin of Entomological Research* 94, 517-524.

Morin P.A., Luikart G. & Wayne R.K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* 19, 208-216.

Nelson J.D., Denisenko O. & Bomsztyk K. (2006). Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nature Protocols* 1, 179-185.

Park P.J. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics* 10, 669-680.

Peterson C.L. & Laniel M.A. (2004). Histones and histone modifications. *Current Biology* 14, R546-R551.

PHA (2008) Draft National Fruit Fly Strategy: March 2008. Plant Health Australia: Canberra, Australia.

Pike N. & Meats A. (2002). Potential for mating between *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy) (Diptera: Tephritidae).*Australian Journal of Entomology* 41, 70-74.

Shelly T.E., McInnis D.O., Pahio E. & Edu J. (2004). Aromatherapy in the Mediterranean fruit fly (Diptera: Tephritidae): sterile males exposed to ginger root oil in prerelease storage boxes display increased mating competitiveness in field-cage trials. *Journal of Economic Entomology* 97, 846-853.

Shelly T.E. (2010). Effects of methyl eugenol and raspberry ketone/cue lure on the sexual behavior of Bactrocera species (Diptera: Tephritidae). *Applied Entomology and Zoology* 45, 349-361.

Thorne A.W., Myers F.A. & Hebbes T.R. (2004). Native chromatin immunoprecipitation. Pp 21-44 in: *Methods in Molecular Biology*, vol. 287: *Epigenetics Protocols* (ed T.O. Tollefsbol). Humana Press Inc., Totowa, N.J.

Tran V., Gan Q. & Chen, X. (2012). Chromatin immunoprecipitation (ChIP) using *Drosophila* tissue. *Journal of Visualized Experiments* 61: 3745.

Wang Y., Yu H., Raphael K. & Gilchrist A.S. (2003). Genetic delineation of sibling species of the pest fruit fly *Bactrocera* (Diptera: Tephritidae) using microsatellites. *Bulletin of Entomological Research* 93, 351-360.

Weldon C.W. (2005). Mass-rearing and sterilisation alter mating behaviour of male Queensland fruit fly, *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae). *Australian Journal of Entomology* 44, 158-163.

Weldon C.W., Prenter J. & Taylor P.W. (2010). Activity patterns of Queensland fruit flies (*Bactrocera tryoni*) are affected by both mass-rearing and sterilization. *Physiological Entomology* 35, 148-53.

Wolda H. (1967). Reproductive isolation between two closely related species of the Queensland fruit fly *Dacus tryoni* (Frogg) and *D. neohumeralis* Hardy (Diptera: Tephritidae) II. Genetic variation in humeral callus pattern in each species as compared with laboratory-bred hybrids. *Australian Journal of Zoology* 15, 515-539.

Zheng W., Peng T., He W. & Zhang H. (2012). High-throughput sequencing to reveal genes involved in reproduction and development in *Bactrocera dorsalis* (Diptera: Tephritidae). *PLoS One* 7, e36463.

Acknowledgements

The authors would like to acknowledge the support of the Molecular Genetics Research Facility and the Australian Genomic Research Facility with genetics component of the project. We would like to thank Mr Michal Lorenc who facilitated the bioinformatics analyses and Mrs Sahana Manoli who facilitated the ChIP-seq preparations.

Appendix I: Draft MS, Chromatin immuno-precipitation

Appendix 1. Draft Manuscript:

Title: Chromatin immuno-precipitation and histone modifications in non-model fruit flies (Diptera: Tephritidae)

Nagalingam Kumaran[1*], Kevin Dudley[1], Michal Lorenc[1], Sahana Manoli[1], Stephen L Cameron[1] and Anthony R Clarke[1, 2]

1. School of Earth, Environmental and Biological Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Qld, Australia

2. Plant Biosecurity Cooperative Research Centre, LPO Box 5012, Bruce, ACT, Australia

*Correspondence

N. Kumaran

School of Earth, Environmental and Biological Sciences,

Queensland University of Technology,

GPO Box 2434, Brisbane, Qld, Australia

k.nagalingam@qut.edu.au

For Journal: Genomics

Word count:

Chromatin immuno-precipitation and histone modifications in non-model fruit flies (Diptera: Tephritidae)

Nagalingam Kumaran, Kevin Dudley, Michal Lorenc, Sahana Manoli, Stephen L Cameron and Anthony R Clarke

Abstract

Modification in histone proteins controls crucial biological processes through regulation of functional gene expressions. Chromatin immuno-precipitation (ChIP) is widely appreciated methodology to study histone proteins, which maps histone modification by precipitating chromatin surrounding DNA with antibodies of interest that can recognize specific histones. While ChIP is extensively used for various cell types successfully in recent times, mapping of histones in non-model species is rarely perceived since the protocol development focused mainly on true model species, which made it difficult to extend the usage of the technique to understand biological processes in non-model organisms. We developed a methodology combining manufacturer's protocol and a protocol followed for *Drosophila* to use for true fruit fly species to study taxa specific evolutionary questions and to assist with management of the global pest species. We elaborate here on the methodology that effectively worked for the major fruit fly species, *Bactrocera dorsalis*, *Ceratitis capitata*, *Zeogodacus cucurbitae* and *Bactrocera tryoni*. Further, we report evidence of epigenetic changes occurring in the genome of tephritid fruit flies for the first time.

Keywords: epigenetics, tephritids, tryoni, dorsalis, cucurbitae, ceratitis, medfly, genome, epigenome, Bactrocera, ChIP-seq, gene expression, H3K, methylation, transcriptome

Introduction

Interactions between DNA and the underlying proteins drive the transcriptional regulation and expression of genes (Felsenfeld & Groudine 2003). Posttranslational modifications in histones for instance play a key role in the functional gene regulation (Peterson & Laniel 2004; Kouzarides 2007). Histone proteins are the chief components of chromatin and influence the transcription directly by altering the packaging of DNA (Berger, 2007). Additional to a mechanism regulating the gene expressions, chromatin and associated histone proteins are involved in DNA repair and DNA replication (Kouzarides 2007). The chromatin states that are found along the genome are known as epigenome, and profiling of the epigenome is crucial to understand the functional gene regulation and expression takes place in any organism (Park 2009).

The technique predominantly used for epigenome profiling is the chromatin-immunoprecipitation (ChIP) followed by identification of DNA fragments through sequencing (ChIP-seq). Briefly, ChIP includes crosslinking the bound protein (DNA) with chromatin, lysing the chromatin extractions and fragmentation of the chromatin (by sonication to generate chromatin fragments of 200-1000 bp size) (Thorne et al. 2004). Then the chromatin is immuno-precipitated through conjugation using specific antibodies that can recognize the bound proteins or protein modifications of interest (Park 2009). Finally, crosslinks are reversed and the free DNA is sequenced to determine the sequences bound by the protein (Nelson et al. 2006).

ChIP has been used in several model species to understand complex biological processes (Furey 2012). However, to the best of our knowledge, there are no ChIP studies on non-model species, and this is perhaps because the methodology developed for model species often fail to work for non-models or at least requires considerable alterations. We tested ChIP methodology described by manufacturers (Abcam) and a methodology followed for the model species *Drosophila* (Tran et al. 2012) in non-model fruit fly species (Diptera: Tephritidae). These methods failed to effectively immuno-precipitate chromatins and resulted in poor recovery because of ineffective cross linking and antibody conjugation. Hence, we modified manufacturer's and *Drosophila* methodologies at crucial steps and found that the revised methodology effectively cross-linked, immunoprecipitated and conjugated antibodies for downstream analyses. We tested this methodology for major tephritid fruit flies, the Oriental fruit fly, *Bactrocera dorsalis* (Hendel), the Mediterranean fruit fly, *Ceratitis capitata* (Weidenmann), melon fly, *Zeogodacus* (*Bactrocera*) *cucurbitae* (Coquillet) and the Queensland fruit fly, *Bactrocera tryoni* (Froggatt). We report here the detailed methodology and the first evidence of histone modifications in genome regions of tephritid fruit flies using antibodies H3K36me3, H3K36me1, H3K27acetylation, H3K27me3, H3K4me2 and H3K4me3.

Tephritid fruit flies are pests of several fruit and vegetable crops (Clarke et al. 2011). Most of the pest species are highly invasive with complex reproductive behaviours (Kumaran et al. 2013; refs). Genetic understanding of developments, behaviours and physiology have been undertaken to understand the genetics behind their reproduction and other physiology (Kumaran et al. 2014a, b and Kumaran & Clarke 2014 refs). As of other organisms, mapping of protein-DNA interactions and epigenetic marks is highly essential for complete understanding of various transcriptional regulations take place in these species to understand evolutionarily important processes to help with sustainable pest management. Hence, this methodology will be of another highly useful resource to understand behaviours of these speciose pest groups.

Methods

Insect source

Bactrocera tryoni was sourced from a colony maintained at the [Queensland] Department of Agriculture and Forestry, Brisbane, Australia. The colony was refreshed annually by introducing wild flies from cultivated fruits. Live flies were killed by chilling and used for the downstream analysis. *Ceratitis capitata* was obtained from a colony maintained at the Department of Agriculture and Food, Perth, Western Australia. *Bactrocera dorsalis* and *Z. cucurbitae* was collected from the infested fruits in Thailand and reared at the plant protection research and development office, Department of Agriculture, Thailand. *Ceratitis capitata*, *B. dorsalis* and *B. cucurbitae* were snap frozen after sexual maturity and stored in RNAlater at -80°C until further analysis.

Reagents

Cross-linking and immuno-precipitation

1. Phosphate buffered saline (DPBS)
2. Protease inhibitor cocktail
3. Formaldehyde
4. Glycine
5. Ribonuclease A
6. Proteinase K
7. Protein A/G beads (ChIP grade)
8. Hering Sperm DNA (HS DNA)
9. Bovine serum albumin (BSA)
10. Antibodies of interest (ChIP grade, Table 1)
11. Sodium Chloride
12. Sodium deoxycholate
13. Sodium bicarbonate
14. Calcium Chloride
15. Sodium Butyrate
16. Tris Hydrogen chloride
17. Triton X-100
18. Ethylene diamine tetra acetic acid (EDTA)
19. Sodium dodecyl sulphate (SDS)

Library preparation

21. Illumina library preparation kit
22. Illumina PCR kit
23. SYBR gold
24. Ampure XP beads

Equipment

1. Bioruptor® UCD-200
2. Water bath
3. Refrigerated centrifuge
4. Rotator
5. Thermocycler
6. Electrophoresis unit and reader
7. UV illuminator for gel excision
8. Illumina Nextseq 500
9. Realtime PCR
10. Bioanalyser
11. Qubit

Procedure

Crosslinking

1. Dissect out tissue (30-40mg, we used heads of 30 male flies per tube) in cold PBS + 1x protease inhibitor (30 µL /ml).

2. Rinse tissue twice with PBS and re-suspend in 200 µL of the same PBS solution with protease inhibitors

3. Cross linking: Add 5.5 µL 37 % formaldehyde. Incubate at RT for 5 minutes, vortex in between

4. Add glycine 0.125M to stop cross linking and incubate at RT for 5 minutes, vortex in between

5. Remove PBS from samples and rinse samples twice in 450 µL of PBS (with protease inhibitor). Samples can be stored at - 20°C now.

Cell lysis

6. Add 200 µL of lysis buffer (protease inhibitor added fresh to buffer) and homogenize using blue homogenizer to get unicellular suspension. Then incubate the lysed tissues at RT for 10 minutes

Shearing

For an effective ChIP, chromatin DNA of 200-300 bp is needed. This will need optimization as different cell lines require different sonication time. The sonication procedure described below shear the chromatin to the size of 200-300 bp for tephritid fruit flies. For a larger size fragments, and for softer tissues sonication time may be reduced.

7. Transfer homogenized tissues to sonication tubes (0.5 ml)

8. Sonicate the lysate to shear chromatin at the following setting: 1.25 to 1.5 h at high settings; 30 sec on: 30 sec off cycle (The setting is for Bioruptor® UCD-200; for a different sonication tool and methods optimization is needed).

9. After sonication, centrifuge at 4°C for 30 sec at 8000 rcf and transfer supernatant to new tube and dilute the sheared chromatin by adding 1 ml of RIPA buffer (protease inhibitor added to RIPA buffer fresh each time)

Input

9a. Aliquot 40 µL of diluted chromatin and add 2 µL of 5M NaCl and incubate at 65°C O/N. This will serve as input for downstream analysis and for quality checks.

9b. Add 2 µL RNAse A and incubate at 65°C for 2 h

9c. Add 5 µL proteinase K incubate at 65°C for 4 h

9d. Purify the DNA using a kit

Immunoprecipitation

Bead preparation

12. 20 µl of protein A/G beads per IP is recommended

13. Wash beads three times in 3X IP dilution buffer

14. Apply to magnet and aspirate dilution buffer off and discard

15. Add 75 ng (per µl beads) of HS DNA + 0.1 µg (per µl beads) of BSA + dilution buffer to twice the bead volume

16. Incubate for 30 min with rotation at RT

17. Wash once with 3X IP dilution buffer

18. Suspend in twice the bead volume of IP buffer

Pre clearing of chromatin

19. Add 20 µl beads to each chromatin sample and incubate at 4°C for 1 hr with rotation

Antibody conjugation

21. Add 5 - 10 µl of antibody of interest to 20 µl beads (for mock add PBS) and add 75 µl PBS

23. Incubate at RT for 1 hr or 4°C for 4 hrs. For effective antibody conjugation to beads perform this step at 4°C

Immunoprecipitation

24. Apply beads with antibody to magnet and discard supernatant

25. Apply beads with sample to magnet and aspirate off supernatant

26. Combine the supernatant of chromatin sample to antibody conjugated beads

27. Incubate O/N at 4°C with rotation

28. Apply sample to magnet and discard supernatant

29. 2X Quick wash with 1 ml of wash buffer for 1 min

30. Wash with 1 ml of wash buffer for 5 min

31. Wash with 1 ml of final wash buffer for 1 min

32. Wash with 1 ml of final wash buffer for 5 min

Elution and Reverse crosslinking

33. Add 120 µl of elution buffer to beads and rotate for 15 min at 30°C (RT)

34. Apply to magnet and transfer supernatant to new tube

35. Add 2 µl (per 40 µl sample) of 5M NaCl and incubate at 60°C O/N

36. Add 2 µl per 40 µl sample of RNAse A and incubate for 2 h

37. Add 5 µl per 40 µl sample of proteinase K and incubate for 4 h

38. Purify the DNA using a kit. This DNA can used for library preparation and sequencing.

Library preparation & Sequencing

We performed library preparation using Illumina library preparation kit and sequenced in Nextseq 500 at the Queensland University of Technology Molecular Genetics Research Facility. We recommend following the platforms and protocols that suit best for the laboratories.

Representative Results

Critical steps for effective ChIP are (i) cross-linking (ii) shearing to the right size (iii) immuno-precipitation and (iv) DNA isolation. Extracted chromatin was checked with agarose gel and Bioanalyzer revealing that the chromatin is cross-linked and shearing (sonication) yielded a right size product for downstream applications (Fig 1).

Antibody conjugation and immunoprecipitation were checked by Qubit assay and the result of conjugation confirmed with Qubit and bioanayzer profiles (Fig 2). Qubit values for chromatin DNA after sonication and DNA concentration after antibody conjugation is presented in table 2. The data shows high quality DNA of ~200 bp size and an effective antibody conjugation (Fib. 2A, B & C).

For sequencing we recommend the protocols that best suit for the laboratories and use the appropriate reagents for library preparation and sequencing. During library preparation several steps need to be optimised to best suit the species. When used Illumina platform libraries on Next-seq 500 we came across several hurdles and modified the procedure slightly (Boxes 1, 2 & 3).

The methodology yielded in successful sequencing of chromatin DNA with high quality reads (Figure 3 & Table 5). After the alignment of reads, we visualized it by uploading to the UCSC browser, which showed enrichment of specific histone proteins (Fig. 4), confirming the methodology has worked well. We were able to see more than 20000 peaks from a pool of 6 samples for H3K36me3 and H3K4me3.

Discussion

Chromatin immunoprecipitation is relatively new and advanced protocol to study the epigenome of an organism that has been mostly used in model systems with almost no known work carried out in non-model systems (Orlando 2000; O'Neill et al. 2006, Collas 2010, Furey 2012 and Gadaleta et al 2015). We developed a protocol to work for a group of non-model true fruit flies by combining a methodology that worked for Drosophila species and a protocol described for confluent animal dishes (manufacturer's protocol). The methods described here worked successfully and yielded high quality chromatin for downstream analyses. We are confident that the protocol will be of highly useful to perform high throughput epigenetic works in tephritid fruit flies and considerably reduce the timing of experiments. In addition, we reported histone modifications in tephritid fruit flies with modifications noticed in regions immuno-precipitated with H3K36me3, H3K36me1, H3K27acetylation, H3K27me3, H3K4me2 and H3K4me3. We are confident that this is the first report on histone modifications in non-

model tephritid fruit flies.

While the methodology was developed by combining two methods, these procedures were modified during critical steps such as cross-linking, shearing and immunoprecipitation. When we followed the *Drosophila* protocol (Tran et al. 2012), cross-linking was unsuccessful for tephritid fruit flies with no yield of chromatin. The more general animal cell methodology failed to cross-link and shear the chromatin to right size perhaps because of more starting material of 1 g tissue (Abcam). The protocol that worked for *Drosophila* also did not work perhaps because of lack of a step in the methodology to stop the cross linking that avoids over cross-linking. Over cross-linking will result in poor chromatin DNA which ultimately affects the immuno-precipitation.

The method we described here used 40 mg of fruit fly heads as starting material as against 1 g suggested by the manufacturer. Further, the Abcam protocol requires 25 ug of chromatin per immunoprecipitation for an effective conjugation because of loss during downstream process; however in the modified protocol we used 240 ng of chromatin as starting material. The use of 40 mg of tissues as starting material yielded chromatin for at least two immuno-precipitations with output DNA of 2 - 20 ng / ul (50 - 500 ng in total) for library preparation depending on the antibodies used. The recommended starting concentration for Illumina library preparation is 5 - 10 ng (Illumina), which suggest that that the methodology is working better for a range of more abundant histone proteins.

Genetic resources such and genome and transcriptome resources are already available for tephritid fruit flies, and already assisting researchers to understand complex biological and behavioural processes (Arthofer et al. 2014; Gilchrist et al. 2014, Gomulski et al. 2012, Spanos et al. 2000, Shen et al. 2011 and Yu et al. 2007). The ChiP methodology described here will add to the already available resources.

Acknowledgement

References

Arthofer, W., Clarke, A. R., Kumaran, N., Prentis, P. J., Schafellner, C., Schlick-Steiner, B. C., ... & Wachter, G. A. (2014). Genomic resources notes accepted 1 December 2013–31 January 2014. *Molecular ecology resources*, *14*(3), 664-665.

Collas, P. (2010). The current state of chromatin immunoprecipitation. *Molecular biotechnology*, *45*(1), 87-100.

Felsenfeld, G., & Groudine, M. (2003). Controlling the double helix. *Nature*, *421*(6921), 448-453

Furey, T. S. (2012). ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nature Reviews Genetics*, *13*(12), 840-852.

Gadaleta, M. C., Iwasaki, O., Noguchi, C., Noma, K. I., & Noguchi, E. (2015). Chromatin immunoprecipitation to detect DNA replication and repair factors. *DNA Replication: Methods and Protocols*, 169-186.

Gilchrist, A. S., Shearman, D. C., Frommer, M., Raphael, K. A., Deshpande, N. P., Wilkins, M. R., ... & Sved, J. A. (2014). The draft genome of the pest tephritid fruit fly Bactrocera tryoni: resources for the genomic analysis of hybridising species. *BMC genomics*, *15*(1), 1153.

Gomulski, L. M., Dimopoulos, G., Xi, Z., Scolari, F., Gabrieli, P., Siciliano, P., ... & Gasperi, G. (2012).

Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, Ceratitis capitata. *PloS one*, *7*(1), e30857.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, *128*(4), 693-705.

Kumaran, N., & Clarke, A. R. (2014). Indirect effects of phytochemicals on offspring performance of Queensland fruit fly, Bactrocera tryoni (Diptera: Tephritidae). *Journal of applied entomology*, *138*(5), 361-367.

Kumaran, N., Balagawi, S., Schutze, M. K., & Clarke, A. R. (2013). Evolution of lure response in tephritid fruit flies: phytochemicals as drivers of sexual selection. *Animal Behaviour*, *85*(4), 781-789.

Kumaran, N., Hayes, R. A., & Clarke, A. R. (2014). Cuelure but not zingerone make the sex pheromone of male Bactrocera tryoni (Tephritidae: Diptera) more attractive to females. *Journal of insect physiology*, *68*, 36-43.

Kumaran, N., Prentis, P. J., Mangalam, K. P., Schutze, M. K., & Clarke, A. R. (2014). Sexual selection in true fruit flies (Diptera: Tephritidae): transcriptome and experimental evidences for phytochemicals increasing male competitive ability. *Molecular ecology*, *23*(18), 4645-4657.

Nelson, J. D., Denisenko, O., & Bomsztyk, K. (2006). Protocol for the fast chromatin immunoprecipitation (ChIP) method. *NATURE PROTOCOLS-ELECTRONIC EDITION-*, *1*(1), 179.

O'Neill, L. P., VerMilyea, M. D., & Turner, B. M. (2006). Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nature genetics*, *38*(7), 835-841.

Orlando, V. (2000). Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends in biochemical sciences*, *25*(3), 99-104.

Park, P. J. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*, *10*(10), 669-680.

Peterson, C. L., & Laniel, M. A. (2004). Histones and histone modifications. *Current Biology*, *14*(14), R546-R551.

Shen, G. M., Dou, W., Niu, J. Z., Jiang, H. B., Yang, W. J., Jia, F. X., ... & Wang, J. J. (2011). Transcriptome analysis of the oriental fruit fly (Bactrocera dorsalis). *PLoS One*, *6*(12), e29127.

Spanos, L., Koutroumbas, G., Kotsyfakis, M., & Louis, C. (2000). The mitochondrial genome of the Mediterranean fruit fly, Ceratitis capitata. *Insect molecular biology*, *9*(2), 139-144.

Thorne, A. W., Myers, F. A., & Hebbes, T. R. (2004). Native chromatin immunoprecipitation. *Epigenetics Protocols*, 21-44.

Tran, V., Gan, Q., & Chen, X. (2012). Chromatin immunoprecipitation (ChIP) using Drosophila tissue. *Journal of visualized experiments: JoVE*, (61).

Yu, D. J., Xu, L., Nardi, F., Li, J. G., & Zhang, R. J. (2007). The complete nucleotide sequence of the mitochondrial genome of the oriental fruit fly, Bactrocera dorsalis (Diptera: Tephritidae). *Gene*, *396*(1), 66-74.

Table 1. Antibodies used for protocol optimization

	Product code (Abcam)	Antibody (ChIP grade)
1	ab9050	Rabbit polyclonal to Histone H3 (tri methyl K36)
2	ab9048	Rabbit polyclonal to Histone H3 (mono methyl K36)
3	ab8580	Rabbit polyclonal to Histone H3 (tri methyl K4)
4	ab7766	Rabbit polyclonal to Histone H3 (di methyl K36)
5	ab6002	Mouse monoclonal to Histone H3 (tri methyl K27)
6	ab4729	Rabbit polyclonal to Histone H3 (acetyl K27)
7	ab1791	Rabbit polyclonal to Histone H3 - Nuclear Loading Control

Table 2. Buffer recipes

Buffer	Recipe
Lysis buffer	50mM Tris-HCl pH 8.0, 1mM $CaCl_2$, 0.2% Triton X-100, 5mM Sodium Butyrate, Protease inhibitor cocktail
RIPA buffer	10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 0.1% SDS, 0.1% Sodium Deoxycholate, 1% Triton X-100, Protease inhibitor cocktail
IP Dilution buffer	1% Triton X-100, 20mM Tris-HCl pH 8.0, 2mM EDTA pH 8.0, 150mM NaCl (add last), Protease inhibitor cocktail
Wash buffer	0.1% SDS, 1% Triton X-100, 20mM Tris-HCl pH 8.0, 2mM EDTA pH 8.0, 150mM NaCl (add last)
Final wash buffer	0.1% SDS, 1% Triton X-100, 20mM Tris-HCl pH 8.0, 2mM EDTA pH 8.0, 500mM NaCl (add last)
Elution buffer	1% SDS, 100mM $NaHCO_3$

Table 3. Recovery of initial chromatin-DNA after cross-linking and output DNA after antibody conjugation for library preparation in tephritid fruit flies

Species	Chromatin DNA after cross-linking (starting material 30 - 40 mg of tissue)	Chromatin DNA after antibody conjugation (starting material based on input DNA concentration = 2.8 to 3.5 µg)
Bactrocera tryoni	248 ng	80.55 ng (H3K36me3)
	244 ng	135.96 ng (H3K36me1)
	302 ng	54.74 ng (H3K4me3)
	272 ng	172.92 ng (H3K4me2)
	340 ng	93.72 ng (H3K27me3)
	332 ng	150.04 ng (H3K27 acetylation)
	250 ng	66.24 ng (Histone 3 nuclear loading control)
	312 ng	< 0.05 (Mock)
Ceratitis capitata	212 ng	73.48 ng (H3K27me3)
	148 ng	40.48 ng (H3K27 acetylation)
	133 ng	51.98 ng (Histone 3 nuclear loading control)
	147 ng	1.76 ng (Mock)
Bactrocera dorsalis	184 ng	78.66 ng (H3K36me1)
	162 ng	64.43 ng (H3K27me3)
	168 ng	1.01 ng (Mock)
Zeogodacus cucurbitae	195 ng	88.42 ng (H3K36me1)
	163 ng	72.95 ng (H3K27me3)
	183 ng	< 0.05 ng (Mock)

Table 3. Recovery of initial chromatin-DNA after cross-linking and output DNA after antibody conjugation for library preparation in tephritid fruit flies

Antibodies	*Bactrocera tryoni*		*Ceratitis capitata*		*Bactrocera dorsalis*		*Zeogodacus cucurbitae*	
	Cross-linked (ng)	Conjugated (ng)	Cross-linked (ng)	Conjugated (ng)	Cross-linked (ng)	Conjugated (ng)	Cross-linked (ng)	Conjugated (ng)
H3K36me3	248	80.55	212	73.48	-	-	-	-
H3K36me1	244	135.96	-	-	184	78.66	195	88.42
H3K4me3	302	54.74	-	-	-	-	-	-
H3K4me2	272	172.92	-	-	-	-	-	-
H3K27me3	340	93.72	-	-	162	64.43	163	72.95
H3K27Acetyl	332	150.04	148	40.48	-	-	-	-
Mock	312	< 0.05	147	1.76	168	1.01	183	< 0.05

Table 4. Output DNA of 200 - 300 bp (enriched) obtained for sequencing from tissues of B. tryoni when using Illumina library preparation kit (starting material 10 ng of antibody conjugated DNA).

Histone antibodies	DNA concentration (Eluted in 40 ul)
H3K36me3	257 - 1440 ng
H3K36me1	808 - 2576 ng
H3K4me3	608 - 1608 ng
H3K4me2	572 - 3288 ng
H3K27me3	528 - 1936 ng
H3K27acetylation	19.5 - 68.2 ng/µl
H3 Nuclear loading control	4.30 - 40.2 ng/µl
Input DNA	10.1 - 43.6 ng/µl

Table 5. Summary sequence data from single illumina run - Total reads, high quality reads etc.

Indexing QC

Lanes	Total Reads	PF Reads	% Reads Identified (PF)	CV	Min	Max
Lane 1	233013360	219931326	97.9531	0.1607	4.951	8.4771
Lane 2	229440604	217221220	97.6945	0.1611	4.9513	8.4623
Lane 3	235203672	222668790	97.8653	0.1606	4.9515	8.4662
Lane 4	230079740	218419190	97.6604	0.1608	4.95	8.4513

Run metrics from all 4 lanes

	Cycles	Yield	Projected Yield	Aligned (%)	Error Rate (%)	Intensity Cycle 1	%≥Q30
Read 1	40	17.13 Gbp	17.13 Gbp	0.98	0.18	8,611	96.91
Read 2	6	2.20 Gbp	2.20 Gbp	0	0	6,137	96.88
Read 3	40	17.10 Gbp	17.10 Gbp	0.96	0.22	9,103	94.61

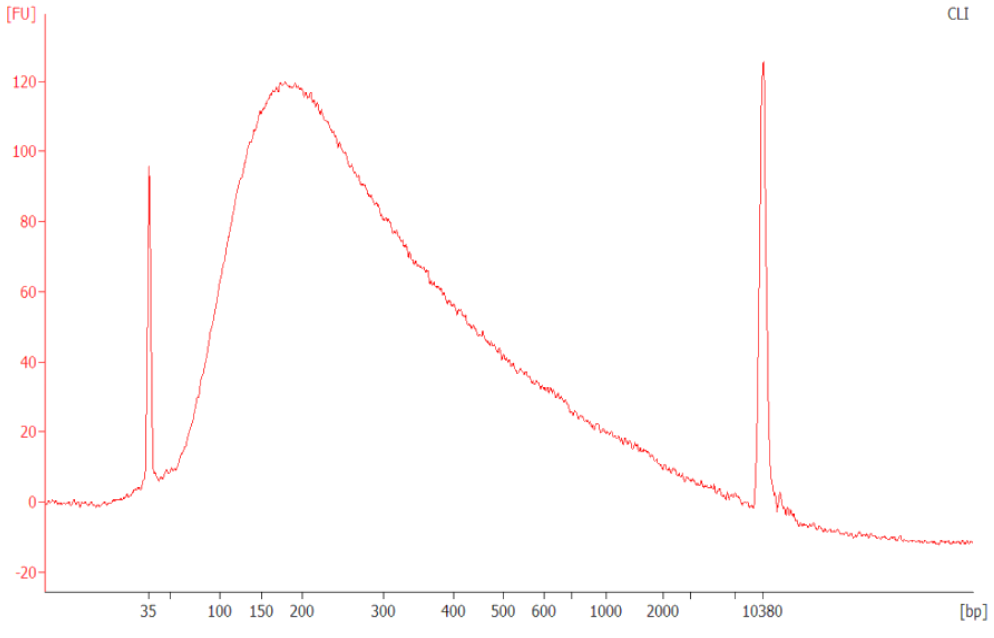


Fig 1. An example of product size after sonication using Bioruptor at 30 sec on, 30 sec off at high speed for 70 - 80 minutes

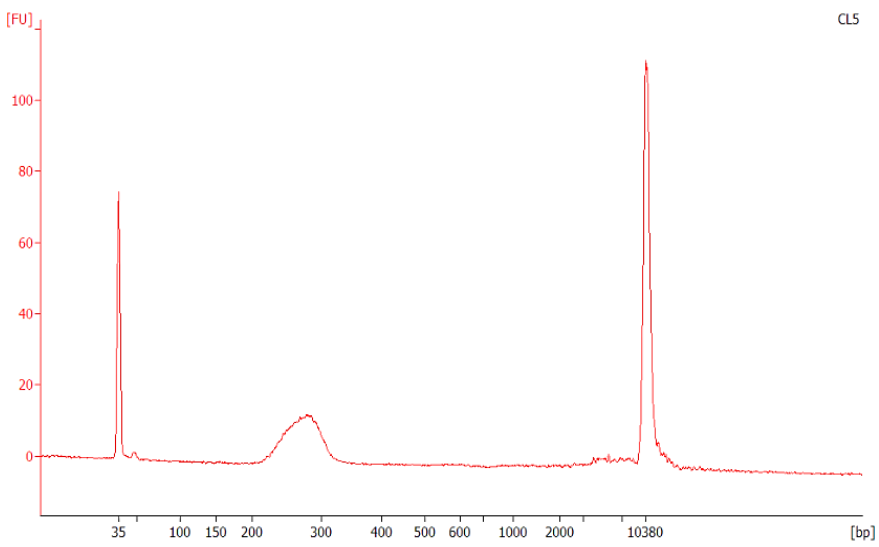


Fig 2A. Bioanalyzer image showing the size and concentration of chromatin DNA immuno-precipitated with H3K27me3 in B. tryoni

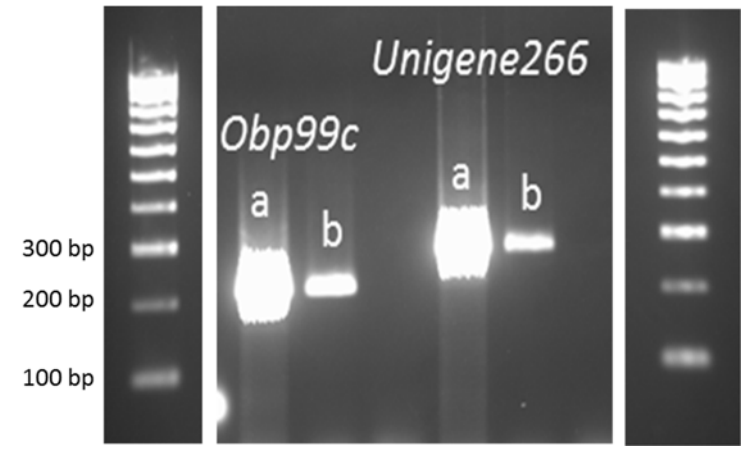


Fig 2B. Evidence of conjugation of H3K27me3 (a) and H3K27 acetylation (b) in *B. tryoni* checked for the genes *Obp 99c* and *Unigene 266*

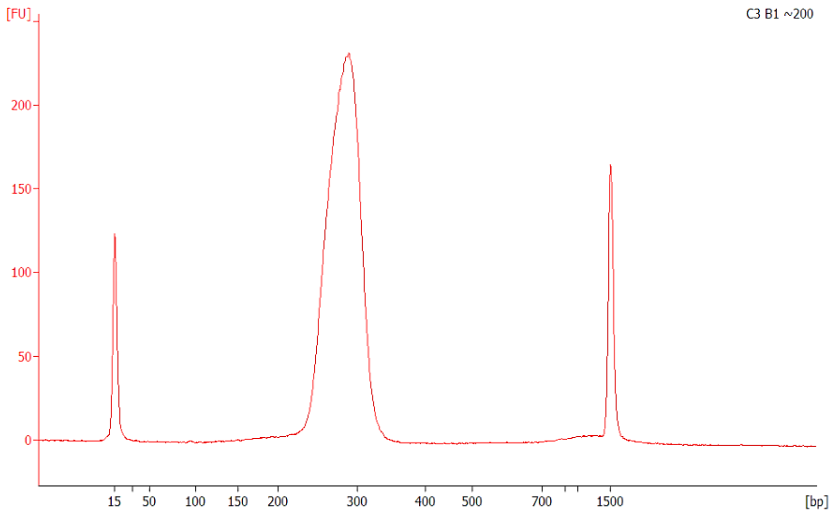


Fig 2C. An example of enriched product size during library preparation size selected for 200-300 bp before sequencing

Figure 3: An example of sequencing outputs from a single illumina run (paired end, 80 cycles) using the library developed with chromatin DNA extracted using the ChIP protocol

Q score distribution

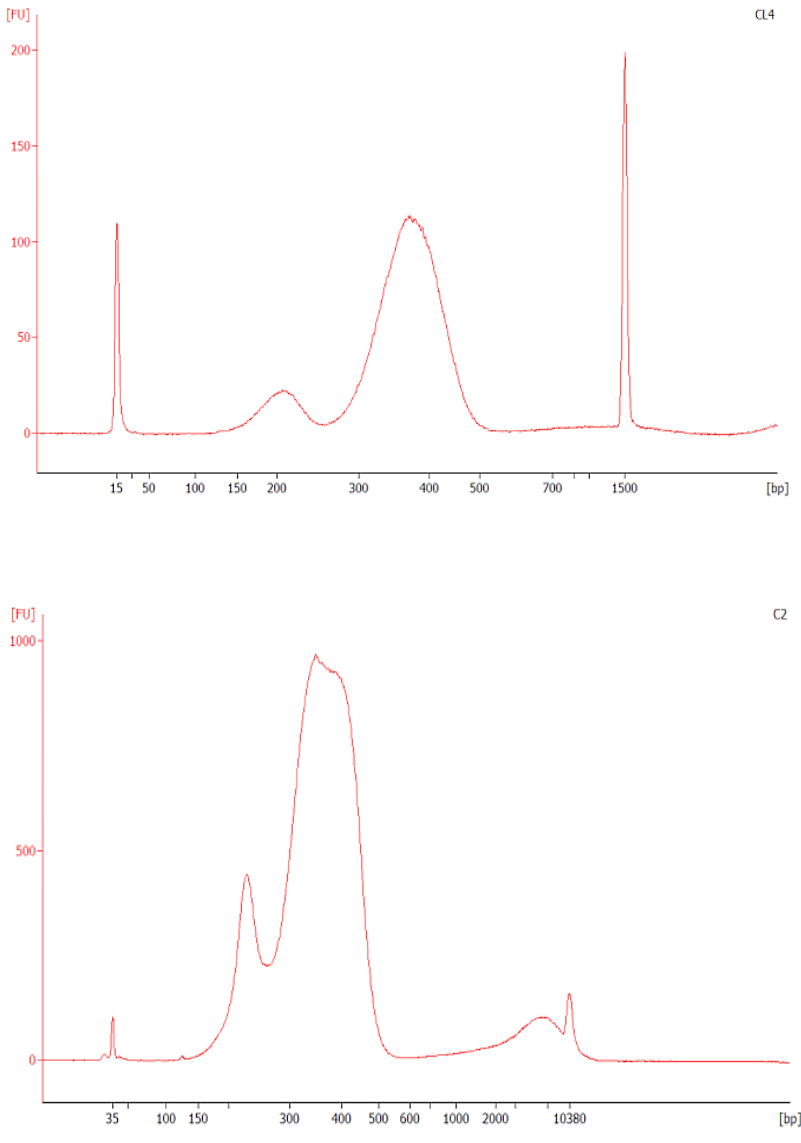
Q score heatmap

Cluster density

Figure 4: An example of ChiP-seq results (snapshots from browsers or tools used for the analyses)

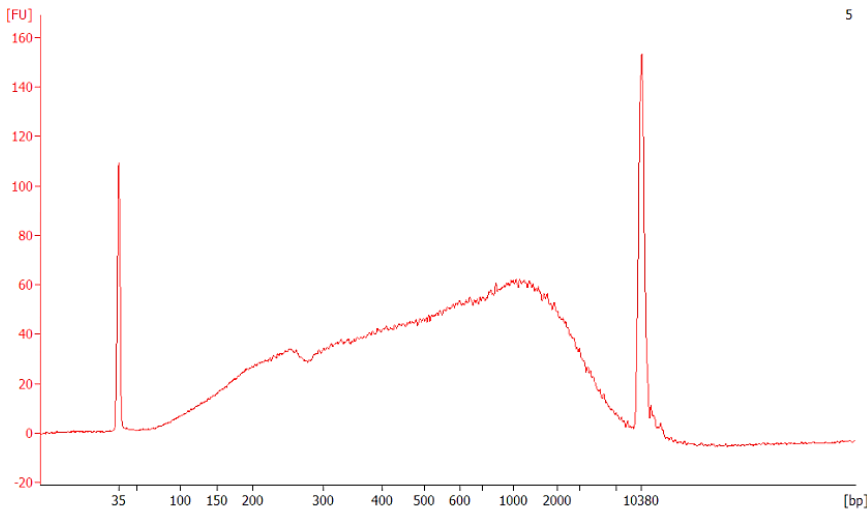
Troubleshoot 1. Daisy chain and over-amplification

We often come across with 'daisy chain' scenario, perhaps due to over amplification during library preparation step. An optimal amplification is necessary to get the preferred size. However, since we size select the product after enrichment at 200 - 300bp, it should not cause any problem when sequencing. If you overly worried about the daisy chain, enrich the product with lesser cycles and gradually improve to optimize the right number of PCR cycles.



Troubleshoot 2. Size selection

Failure of antibody conjugation is primarily because of wrong size of the product. So size selection is critical for effective recovery and successful immunoprecipitation. Always check your product size before proceeding to the conjugation step. Shearing of chromatin DNA to required size of 200-300 bp needs sonicating samples for 70 to 80 minutes if you using Bioruptor. During sonication, make sure to topup the water tank with ice. Shearing by keeping the produce in cold water rather than ice yield poor shearing. In between changing the ice and water, keep the samples on ice. After sonication, make sure the centrifuge is ready at 4C for spinning. It need to be optimized if you using other sonication methods. However, if you suing soft tissues and more high power sonicators perhaps reduced timing should be fine. Since we used the whole head of the flies which is sclerotized, softer tissues might need lesser sonication. But for the most part of the flies, the time or 70 - 80 minutes could be the maximum.



The figure shows poor shearing when followed the protocol developed for model species

Appendix II: Draft MS, Female transcription factors

Appendix 2. Draft Manuscript

Title: Intrinsic female transcription factors regulating mating and post mating physiology

Nagalingam Kumaran[1*], Peter J Prentis[1], Wan Der Burg Chloe[1], Xin-Ji[2] and Anthony R Clarke[1, 3]

3. School of Earth, Environmental and Biological Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Qld, Australia
4.
5. Plant Biosecurity Cooperative Research Centre, LPO Box 5012, Bruce, ACT, Australia

*Correspondence

N. Kumaran

School of Earth, Environmental and Biological Sciences,

Queensland University of Technology,

GPO Box 2434, Brisbane, Qld, Australia

k.nagalingam@qut.edu.au

For Journal: Genome Biology and Evolution

Word count:

Intrinsic female transcription factors regulating mating and post mating physiology

Nagalingam Kumaran[1*], Peter J Prentis[1], Wan Der Burg Chloe[1], Xin-Ji[2] and Anthony R Clarke[1, 3]

Abstract

Female mating and remating to some extent is mediated by male genotype and environment through indirect genetic effects (IGEs) in females. While there is an extensive theoretical and empirical work on post-mating physiological modulation in females through IGEs, female regulating factors of mating and re-mating remain largely unresolved albeit females being a decisive gender to choose whether to mate or not, and with whom in most of the species. In tephritid fruit fly species *Bactrocera tryoni*, females mated with males fed on secondary plant compounds become unreceptive for re-mating when compared with females mated with normal unfed males. We did comparative transcriptome analyses with virgin mature females, females mated with normal unfed males and females mated with males fed on plant compounds to profile the transcriptional factors regulating mating (virgin vs normal mated) and remating (normal mated vs mated with fed males). There were several gene transcripts and Gene Ontology terms over or under represented in all comparison categories. We discuss the up regulated and down regulated genes and GO categories with respect to their role in regulating mating, remating and other post mating physiological changes.

Key words: polyandry, transcriptome, sexual selection, sterile insect technique, cuelure, methyl eugenol, zingerone, tephritidae, fruit fly, indirect genetic effects

Introduction

While single or a few matings is sufficient for females to maximize their reproductive success, polyandry and repeated mating is ubiquitous (Arnqvist & Nilsson 2000). Polyandry help females chose good genes, sperm or genetic materials by selecting high quality males (Fedorka & Mousseau 2002). However, a pile of empirical evidences on the other hand suggest mating is extensively costly and multiple mating incurs huge cost physiologically to invest in the copulation and behaviorally so to mate selection and avoid predation risks (Baer & Schmid-Hempel 2001; Colegrave et al. 2002). Mechanisms mediating re-mating (polyandry) also increase further in complexity as the factors governing re-mating are potentially controlled by three sources: male factor, female factor and the environment. Regulation of re-mating by male genotype and environment through indirect genetic effects (IGEs) in females is well received as there is an extensive theoretical and empirical work on post-mating physiological modulation in females by male accessory gland proteins and sperms (Hosken et al. 2002; Wolf et al. 1998; Gillott 2003). Surprisingly, female regulating factors of mating and re-mating remain largely unresolved albeit females being a decisive gender to choose whether to mate or not, and with whom in most of the species (Andersson & Simmons 2006).

In true fruit flies (Diptera: Tephritidae), males are polygynous, and it has long been assumed that females are monandrous. However, recent empirical evidences suggest polyandry is prevalent in females (Whittier & Shelly 1993; Opp & Prokopy 2000; Song et al. 2007). In these flies, male mating success is mediated by secondary plant compounds such as methyl eugenol, raspberry ketone and zingerone that are not part of fly's diets (Shelly 2010; Kumaran et al. 2013). These compounds are traced in host and non-host plants and males show very strong olfactory and gustatory response to these compounds. It is evidenced that the plant compounds are mediating sexual selection in *Bactrocera* flies, and that males fed on the plant compounds are selected by females (Kumaran et al. 2013). Further analyses confirmed that the mating success is mediated by males releasing sexier pheromones which subsequently attracting females, and that males becoming physically active after feeding on plant compounds that is an additional factor to achieve greater mating (Kumaran et al. 2014a, b).

The secondary plant compounds exerted significant physiological changes in females through males after females mated with fed males. In a study using species *Bactrocera tryoni* and *B. dorsalis*, these compounds altered post mating physiology of females, which being increased egg production and changes in re-mating pattern when females mated with fed males (Kumaran et al. 2013). For instance, females became unreceptive for further matings when first mated with fed males in *B. tryoni*, and fecundity was greater in *B. dorsalis* and *B. tryoni* when mated with fed males (Kumaran et al. 2013). While mechanisms regulating male mating success were thoroughly tested, female factors were largely ignored in *Bactrocera* and other systems assuming that female post mating physiology is largely driven by males with which females mate. However, female genetic makeup mediates to some extent to regulate mating which is particularly possible give that females being choosy of mating and re-mating in majority of the cases (Andersson & Simmons 2006).

Transcriptome profiling of females to understand their reproduction has been studied and is helpful understand complex behaviours and mechanism mediating mating. For instance, stage specifc transcriptome profiling has been done in many species that explained factors regulating reproduction (Gomulski et al. 2012; Zheng et al. 2012). A very few studies on female specific transcriptional regulation of development and mating has been studied previously in tephritid fruit flies and other systems (McGraw et al. 2008; Gonclaves et al. 2013). Although relatively few genes were found to be differentially regulated in females post mating, the results provided insights on female development and

mating in *C. capitata* and *D. melanogaster*.

Hence, in this study we compared virgin and mated females to understand transcriptional factors regulating mating, but most importantly, we compared females mated with normal males and males fed on plant compounds to understand what causes females to re-mate using RNA-seq analyses. Any difference in females mated with normal males compared with fed males should inform us factors regulating re-mating as females mated with fed males would be showing reduced re-mating tendency.

Methods

Insect source

Bactrocera tryoni were obtained as pupae from the rearing facility at the [Queensland Government] Department of Agriculture Fisheries and Forestry, Brisbane. Emerged flies were provided with water, sugar and protein hydrolysate ad libitum and maintained at 27 °C and 70% RH in a room illuminated with natural light in addition to fluorescent lighting between 07:00 and 16:00 hours every day.

Experimental groups

Flies were sexed within 2 days of emergence, and males and females housed separately for analyses. Totally three groups of females were maintained: mature virgin females, females mated with normal males and females mated with males fed on plant compound zingerone. To obtain lure-fed males, flies were provided with 1.5 mL of zingerone (10 lg/ lL of 95% ethanol, Sigma-Aldrich, CHEME, GmbH, Germany, >96% purity) on a cotton wick placed on inverted petri dish for 2 h. The concentrations, dilutions and presentation methods are based on previous studies on other Bactrocera flies (Shelly & Villalobos 1995; Hee & Tan 1998). To obtain mated females, 50 females were housed with 50 normal or fed males in small cages (30 x 15 x 15 cm). Totally two cages per group (virgin females, females mated with fed male and females mated with unfed male) were maintained.

RNA isolation, cDNA library synthesis, EST assembly and annotation

A total of 20 females (14 days old) from each of three groups (virgin females, females mated with fed male and females mated with unfed male), were collected and snap-frozen in liquid nitrogen next morning after flies mated with respective males. Total RNA was extracted from whole bodies using Trizol and purified with a Qiagen RNeasy kit following the manufacturer's instructions. Detailed methodology on isolation, library construction, assembly and annotation, along with sequence files, contigs, unigenes and sequence assembly are presented in the companion study Arthofer et al. (2014).

Analysis of differentially expressed genes

Sequencing reads were mapped to unigenes from two samples and converted to Fragments Per kb per Million fragments (FPKM) to determine which genes were differentially expressed (DEGs) between lure-fed and unfed conditions. P value threshold for the analysis was determined using false discovery rate (FDR). Smaller FDR value and larger FPKM ratio shows a greater difference in expression levels. In this analysis, we chose FDR ≤ 0.001 and FPKM ratio larger than two for significant expression abundance between lure-fed and unfed flies. We undertook gene set enrichment analysis to determine whether particular GO categories were overrepresented in the DEGs. For the identification of significantly enriched metabolic pathways or signal transduction pathways, we mapped all DEGs to terms to the KEGG database using enzyme codes. Pathways with Q value ≤ 0.05 were considered significantly enriched in DEGs.

Results

Transcriptome summary

DEGs virgin vs normal mated females (Fig 1a & b)

DEGs mated with fed-male vs normal male

There were 89 genes differentially expressed in females mated with males fed with lures with 70 genes upregulated and 19 genes down regulated (Fig 2a & b; Table 1 & 2). Investigation of gene categories revealed enrichment of several GO terms ($p < 0.05$). There were 126 GO categories enriched within biological processes GO term with purine metabolic and biosynthetic processes, ribonucleoside metabolic processes, nucleoside metabolic processes, cellular metabolic compound salvage and glycosyl compound metabolic processes being the most significantly enriched GO functions (Table 3a). Within cellular processes GO term, six GO terms showed significant enrichment with the most significant being nucleosome, DNA bending and DNA packing complex (Table 3b). There were 37 GO categories enriched within molecular function GO category with transferase activity, transferring pentosyl groups, S-methyl-5-thioadenosine phosphorylase activity, purine nucleobase binding, purine-nucleoside phosphorylase activity, phosphate ion binding and nucleobase binding function most significantly enriched among other GO functions (Table 3c).

There were 13 GO categories significantly down regulated when females mated with lure-fed males, containing nine biological processes term, three cellular component term and one molecular function term (Table 4). While the GO terms downregulated are lesser compared to up regulated GO terms, the transcripts under each of the down regulated categories perhaps regulate female re-mating as females after mating with lure-fed males become unreceptive for subsequent mating in this species.

DEGs virgin vs females mated with lure-fed males

There were 158 genes differentially expressed in females mated lure-fed males compared with virgin females with 124 genes upregulated and 34 genes down regulated (Fig 3a & b). There were several GO terms ($p < 0.05$) enriched with 113 categories within biological processes GO term with entry into host cell, movement in host environment, symbiotic interaction, nucleic acid-protein covalent cross-linking and RNA-protein covalent cross-linking being the most significantly enriched GO functions (Table 5a). Within cellular processes GO term, 20 GO terms showed significant enrichment with the most significant being host intracellular part, host organelle outer membrane, host cell cytoplasmic vesicle membrane and host cell mitochondrial membrane (Table 5b). There were six GO categories enriched within molecular function GO category with structural constituent of cuticle, structural molecule activity and RNA-directed RNA polymerase activity (Table 5c). There were 37 GO terms down regulated in females mated with lure-fed males when compared with virgin females with regulation of cellular metabolic process, regulation of macromolecule metabolic process, regulation of primary metabolic process, signal transduction, negative regulation of biological process and regulation of metabolic process being the most significantly down regulated GO terms (Table 6).

Discussion

Results summary

Genes regulating mating (virgin vs normal mated)

Post mating physiology (virgin vs normal mated)

Re-mating (normal mated vs lure-fed mated)

Taxa specific transcripts and short peptides

Most of transcripts are extremely specific to taxa or genus with unknown functions. The reason perhaps is that female specific transcriptome resources in general are lacking for any organisms.

Implications for SIT

Acknowledgement

References

Andersson, M., & Simmons, L. W. (2006). Sexual selection and mate choice. *Trends in Ecology & Evolution*, *21*(6), 296-302.

Arnqvist, G., & Nilsson, T. (2000). The evolution of polyandry: multiple mating and female fitness in insects. *Animal behaviour*, *60*(2), 145-164.

Arthofer, W., Clarke, A. R., Kumaran, N., Prentis, P. J., Schafellner, C., Schlick-Steiner, B. C., ... & Wachter, G. A. (2014). Genomic resources notes accepted 1 December 2013–31 January 2014. *Molecular ecology resources*,*14*(3), 664-665.

Baer, B., & Schmid-Hempel, P. (2001). Unexpected consequences of polyandry for parasitism and fitness in the bumblebee, Bombus terrestris.*Evolution*, *55*(8), 1639-1643.

Colegrave, N., Kotiaho, J. S., & Tomkins, J. L. (2002). Mate choice or polyandry: reconciling genetic compatibility and good genes sexual selection.*Evolutionary Ecology Research*, *4*(6), 911-917.

Fedorka, K. M., & Mousseau, T. A. (2002). Material and genetic benefits of female multiple mating and polyandry. *Animal Behaviour*, *64*(3), 361-367.

Gillott, C. (2003). Male accessory gland secretions: modulators of female reproductive physiology and behavior. *Annual Review of Entomology*, *48*(1), 163-184.

Gomulski, L. M., Dimopoulos, G., Xi, Z., Scolari, F., Gabrieli, P., Siciliano, P., ... & Gasperi, G. (2012). Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, Ceratitis capitata. *PloS one*, *7*(1), e30857.

Gonçalves, V. R., Sobrinho, I. S., Malagó, W., Henrique-Silva, F., & Brito, R. A. (2013). Transcriptome analysis of female reproductive tissues of Anastrepha obliqua and molecular evolution of eggshell proteins in the fraterculus group. *Insect molecular biology*, *22*(5), 551-561.

Hosken, D. J., Uhia, E., & Ward, P. I. (2002). The function of female accessory reproductive gland secretion and a cost to polyandry in the yellow dung fly. *Physiological Entomology*, *27*(2), 87-91.

Kumaran, N., Balagawi, S., Schutze, M. K., & Clarke, A. R. (2013). Evolution of lure response in tephritid fruit flies: phytochemicals as drivers of sexual selection. *Animal Behaviour*, *85*(4), 781-789.

Kumaran, N., Hayes, R. A., & Clarke, A. R. (2014). Cuelure but not zingerone make the sex pheromone

of male Bactrocera tryoni (Tephritidae: Diptera) more attractive to females. *Journal of insect physiology*, *68*, 36-43.

Kumaran, N., Prentis, P. J., Mangalam, K. P., Schutze, M. K., & Clarke, A. R. (2014a). Sexual selection in true fruit flies (Diptera: Tephritidae): transcriptome and experimental evidences for phytochemicals increasing male competitive ability. *Molecular ecology*, *23*(18), 4645-4657.

McGraw, L. A., Clark, A. G., & Wolfner, M. F. (2008). Post-mating gene expression profiles of female Drosophila melanogaster in response to time and to four male accessory gland proteins. *Genetics*, *179*(3), 1395-1408.

Opp, S. B., & Prokopy, R. J. (2000). Multiple mating and reproductive success of male and female apple maggot flies, Rhagoletis pomonella (Diptera: Tephritidae). *Journal of insect behavior*, *13*(6), 901-914.

Shelly, T. (2010). Effects of methyl eugenol and raspberry ketone/cue lure on the sexual behavior of Bactrocera species (Diptera: Tephritidae). *Applied Entomology and Zoology*, *45*(3), 349-361.

Song, S. D., Drew, R. A., & Hughes, J. M. (2007). Multiple paternity in a natural population of a wild tobacco fly, Bactrocera cacuminata (Diptera: Tephritidae), assessed by microsatellite DNA markers. *Molecular Ecology*,*16*(11), 2353-2361.

Whittier, T. S., & Shelly, T. E. (1993). Productivity of singly vs. multiply mated female Mediterranean fruit flies, Ceratitis capitata (Diptera: Tephritidae). *Journal of the Kansas Entomological Society*, 200-209.

Wolf, J. B., Brodie III, E. D., Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998). Evolutionary consequences of indirect genetic effects. *Trends in ecology & evolution*, *13*(2), 64-69.

Zheng, W., Peng, T., He, W., & Zhang, H. (2012). High-throughput sequencing to reveal genes involved in reproduction and development in Bactrocera dorsalis (Diptera: Tephritidae). *PLoS One*, *7*(5), e36463.

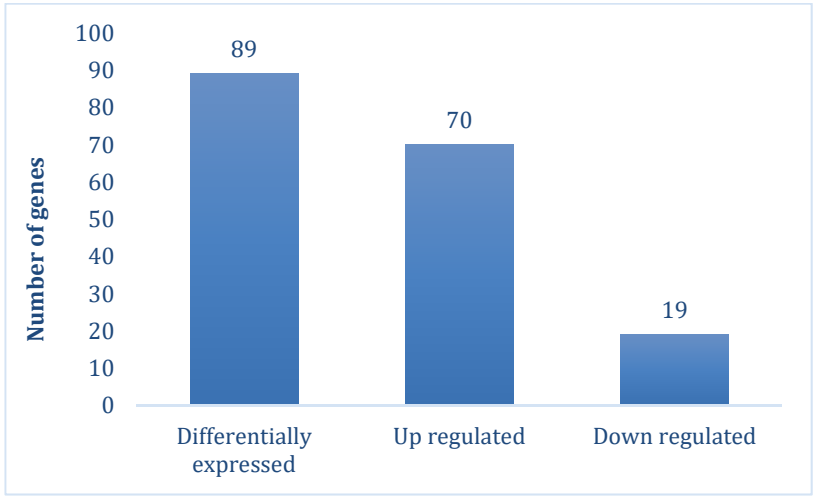


Figure. 3a. Differentially expressed genes in females mated with lure-fed males. Expression profiles compared with females mated with normal unfed males

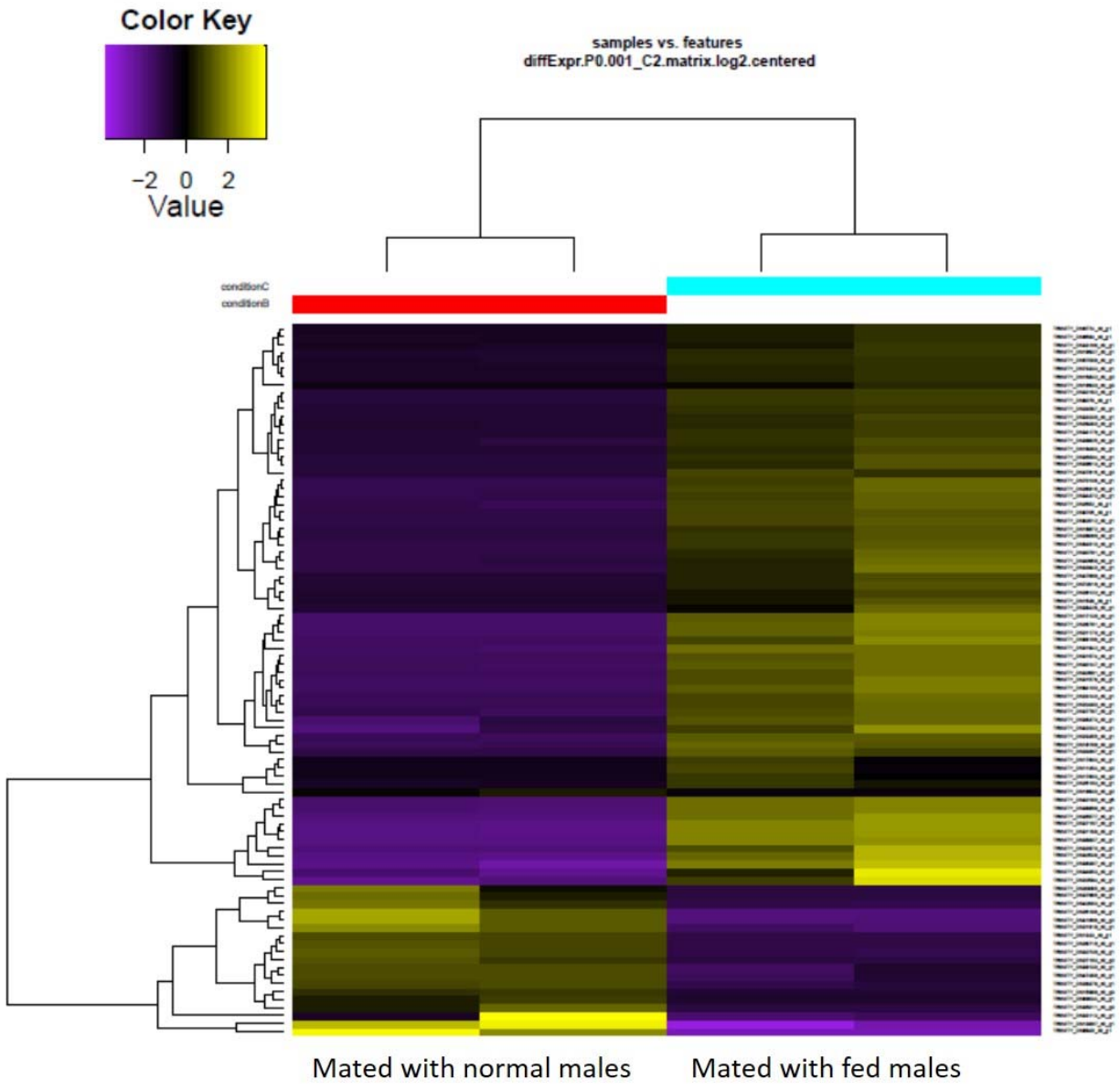


Figure 3b. Heatmap of transcriptome profile showing differential expression of gene transcripts in females mated with normal and lure-fed males

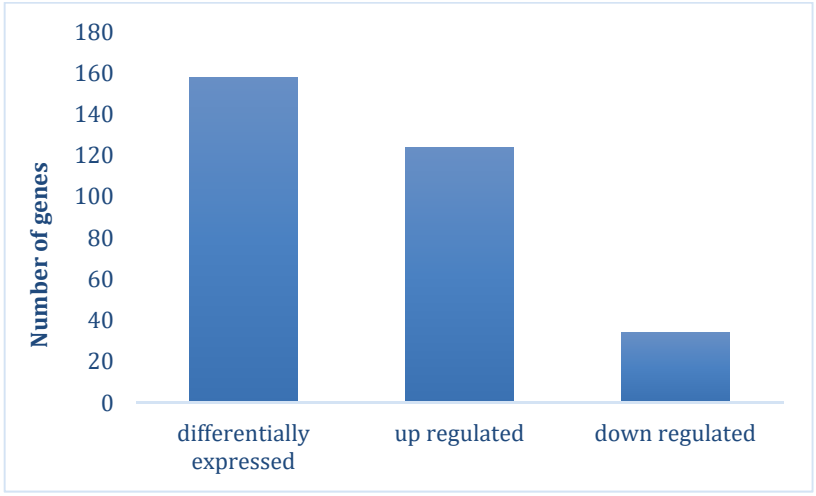


Figure 2a. DEGs mated with fed-male vs virgin

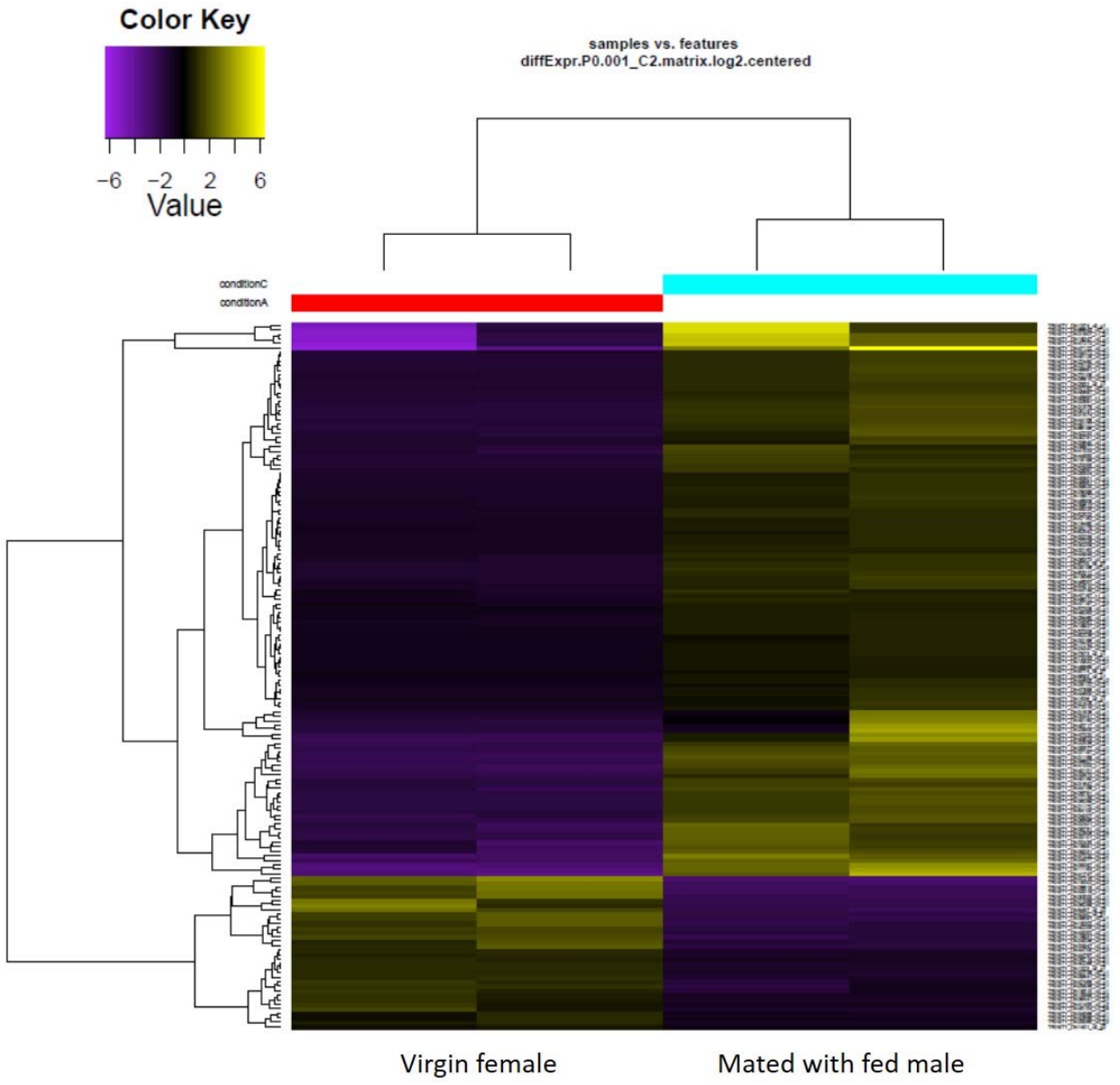


Figure 2b. Heatmap transcriptome profile showing differential expression of gene transcripts in virgin females and females mated with lure-fed males

Table 1. Up regulated genes and functions

Gene ID	P	FDR	Functions	Reference sp
TRINITY_DN36607_c0_g1	3.24E-27	2.73E-22	unknown	
TRINITY_DN35577_c0_g1	7.10E-24	2.99E-19	unknown	c capitata
TRINITY_DN18198_c0_g2	6.79E-21	1.91E-16		B dorsalis
TRINITY_DN31168_c0_g1	2.07E-20	4.37E-16	unknown	
TRINITY_DN36898_c0_g1	2.71E-20	4.56E-16	unknown	B dorsalis
TRINITY_DN26791_c0_g1	1.80E-18	2.16E-14	unknown	c capitata
TRINITY_DN23355_c0_g1	1.88E-17	1.76E-13		B dorsalis; D melano
TRINITY_DN39613_c0_g1	3.84E-17	3.21E-13	unknown	B dorsalis
TRINITY_DN38387_c0_g1	4.19E-17	3.21E-13		B dorsalis; D melano
TRINITY_DN37197_c0_g1	5.34E-17	3.75E-13	unknown	
TRINITY_DN23140_c0_g1	8.24E-17	5.34E-13	unknown	
TRINITY_DN31642_c0_g1	2.79E-16	1.57E-12	unknown	B dorsalis
TRINITY_DN32536_c0_g1	3.97E-16	2.09E-12	unknown	B dorsalis
TRINITY_DN31576_c0_g1	8.78E-16	4.35E-12	unknown	
TRINITY_DN30147_c0_g1	1.38E-15	6.44E-12	unknown	B dorsalis
TRINITY_DN32691_c0_g1	1.00E-14	4.21E-11	unknown	B dorsalis
TRINITY_DN31574_c0_g1	1.91E-14	7.65E-11	unknown	
TRINITY_DN33970_c0_g1	1.12E-13	4.29E-10	unknown	
TRINITY_DN26915_c0_g1	7.10E-13	2.60E-09	unknown	
TRINITY_DN64100_c0_g1	1.51E-12	5.14E-09		D melano;
TRINITY_DN17139_c0_g1	1.52E-12	5.14E-09	unknown	
TRINITY_DN33239_c0_g1	6.09E-12	1.98E-08		B dorsalis
TRINITY_DN40180_c0_g5	1.05E-11	3.27E-08	Sarcotoxin; Cecropin	Sarcophaga peregrina; Hyalophora cecropia
TRINITY_DN35504_c0_g1	2.06E-11	6.21E-08	unknown	B dorsalis
TRINITY_DN18460_c0_g1	3.82E-11	1.11E-07	unknown	B dorsalis
TRINITY_DN37787_c0_g1	6.44E-11	1.81E-07	unknown	B dorsalis
TRINITY_DN66196_c0_g1	7.36E-11	1.94E-07	unknown	
TRINITY_DN19930_c0_g3	7.37E-11	1.94E-07		D melano; B dorsalis
TRINITY_DN21170_c0_g1	8.19E-11	2.09E-07	unknown	
TRINITY_DN32162_c0_g1	2.08E-10	5.16E-07		B dorsalis; Human
TRINITY_DN22480_c0_g1	4.55E-10	1.10E-06	unknown	B dorsalis
TRINITY_DN28099_c0_g1	6.58E-10	1.54E-06	unknown	B dorsalis
TRINITY_DN39825_c0_g2	7.68E-10	1.70E-06	Cytosol aminopeptidase	c capitata;
TRINITY_DN33640_c0_g1	7.88E-10	1.70E-06	unknown	
TRINITY_DN19937_c0_g1	1.06E-09	2.24E-06	unknown	B dorsalis
TRINITY_DN72019_c0_g1	1.62E-09	3.24E-06	unknown	B dorsalis
TRINITY_DN19930_c0_g2	1.96E-09	3.75E-06		D melano; B dorsalis
TRINITY_DN37815_c0_g2	2.73E-09	5.11E-06	unknown	B dorsalis
TRINITY_DN24697_c0_g1	2.79E-09	5.12E-06	unknown	B dorsalis

TRINITY_DN25193_c0_g1	3.39E-09	6.07E-06		D melano; C capitata
TRINITY_DN18872_c0_g1	3.92E-09	6.88E-06	unknown	c capitata
TRINITY_DN2552_c0_g1	9.11E-09	1.45E-05	unknown	B dorsalis
TRINITY_DN22584_c0_g1	9.11E-09	1.45E-05		B dorsalis
TRINITY_DN36373_c0_g1	1.05E-08	1.64E-05	Reverse transcriptase	D melano
TRINITY_DN6705_c0_g1	1.73E-08	2.60E-05		C capitata
TRINITY_DN36426_c0_g1	1.76E-08	2.60E-05	unknown	Danaus plexippus
TRINITY_DN17803_c0_g1	2.20E-08	3.20E-05	unknown	
TRINITY_DN62012_c0_g1	2.32E-08	3.31E-05	unknown	
TRINITY_DN30781_c0_g1	2.75E-08	3.80E-05	unknown	B dorsalis
TRINITY_DN54510_c0_g1	4.20E-08	5.71E-05	unknown	
TRINITY_DN34653_c0_g1	4.28E-08	5.73E-05	unknown	
TRINITY_DN17803_c0_g2	5.90E-08	7.65E-05	unknown	B dorsalis; C capitata
TRINITY_DN28122_c0_g1	7.88E-08	0.000101	linker histone H1 and H5 family	B dorsalis; Chironomous thummi thummi
TRINITY_DN34372_c0_g1	8.83E-08	0.000111	unknown	B dorsalis
TRINITY_DN11354_c0_g2	9.37E-08	0.000116	unknown	
TRINITY_DN9774_c0_g1	1.33E-07	0.000159	unknown	B dorsalis
TRINITY_DN42232_c0_g1	2.24E-07	0.000263	short chain dehydrogenase; Enoyl-(Acyl carrier protein) reductase; Farnesol dehydrogenase	B dorsalis, A. agepti
TRINITY_DN70106_c0_g1	2.96E-07	0.000342	unknown	C capitata
TRINITY_DN29460_c0_g1	3.12E-07	0.000355	unknown	B dorsalis
TRINITY_DN9593_c0_g1	3.77E-07	0.000418	unknown	B dorsalis
TRINITY_DN34179_c0_g1	5.93E-07	0.000641	Sperm outer dense fibre protein	B dorsalis; Danio rerio
TRINITY_DN33196_c0_g1	6.33E-07	0.000667	unknown	C capitata
TRINITY_DN47898_c0_g1	6.41E-07	0.000667	unknown	B dorsalis
TRINITY_DN24567_c0_g1	6.60E-07	0.000677	unknown	B dorsalis
TRINITY_DN6076_c0_g1	6.66E-07	0.000677	unknown	
TRINITY_DN30958_c0_g1	7.15E-07	0.000712	unknown	B dorsalis
TRINITY_DN1536_c0_g1	7.18E-07	0.000712	Testis expressed	B dorsalis; Mouse
TRINITY_DN15832_c0_g2	7.72E-07	0.000757	Lectin C-type domain	B dorsalis; Ovis (sheep)
TRINITY_DN57008_c0_g1	8.68E-07	0.000841	unknown	B dorsalis
TRINITY_DN73333_c0_g1	9.78E-07	0.000926	unknown	B dorsalis

Table 2. Down regulated genes and functions

Gene ID	P	FDR	Functions	Reference sp
TRINITY_DN13697_c0_g1	1.12E-19	1.57E-15	unknown	
TRINITY_DN8949_c0_g1	2.13E-18	2.24E-14	unknown	
TRINITY_DN41896_c0_g2	2.16E-16	1.30E-12	unknown	Reductase like protein
TRINITY_DN25186_c0_g1	4.16E-15	1.85E-11	binding nuclear protein	D melano; A mellifera
TRINITY_DN20065_c0_g2	7.79E-10	1.70E-06	unknown	
TRINITY_DN27193_c0_g2	1.49E-09	3.07E-06	unknown	
TRINITY_DN15988_c0_g3	1.83E-09	3.60E-06		B dorsalis;
TRINITY_DN21918_c0_g1	4.31E-09	7.42E-06		B dorsalis
TRINITY_DN32739_c0_g1	4.41E-09	7.43E-06	unknown	
TRINITY_DN35011_c0_g3	7.68E-09	1.27E-05	nuclear transport factor	D melano; C capitata
TRINITY_DN26719_c0_g1	1.16E-08	1.79E-05	unknown	
TRINITY_DN42003_c0_g1	2.50E-08	3.51E-05	unknown	
TRINITY_DN1323_c0_g1	4.93E-08	6.49E-05	unknown	
TRINITY_DN37965_c0_g1	1.15E-07	0.00014	Ribosomal protein	B dorsalis; D melano
TRINITY_DN28376_c0_g1	1.34E-07	0.000159	Transposable element Tc1 transposase	B dorsalis; C elegans
TRINITY_DN69034_c0_g1	3.70E-07	0.000416	Occludin homology domain	B dorsalis; D melano
TRINITY_DN39130_c0_g1	4.68E-07	0.000512	Ob56A - Odorant binding protein	B dorsalis; D mealno
TRINITY_DN47468_c0_g1	6.13E-07	0.000654	unknown	
TRINITY_DN33113_c0_g1	9.34E-07	0.000895	unknown	

Table 3a. GO term up regulated in the category Biological processes (mated with fed-male vs normal mal)

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0043101	purine-containing compound salvage	15	2	0.00013
GO:0046128	purine ribonucleoside metabolic process	120	3	0.000338
GO:0042278	purine nucleoside metabolic process	126	3	0.000389
GO:0072522	purine-containing compound biosynthetic process	128	3	0.000416
GO:0009119	ribonucleoside metabolic process	142	3	0.000558
GO:0009116	nucleoside metabolic process	156	3	0.00073
GO:0043094	cellular metabolic compound salvage	38	2	0.000847
GO:1901657	glycosyl compound metabolic process	170	3	0.000932
GO:0019523	L-idonate metabolic process	1	1	0.001442
GO:0046176	aldonic acid catabolic process	1	1	0.001442
GO:0046183	L-idonate catabolic process	1	1	0.001442
GO:0072521	purine-containing compound metabolic process	224	3	0.00214
GO:0006738	nicotinamide riboside catabolic process	2	1	0.002257
GO:0046495	nicotinamide riboside metabolic process	2	1	0.002257
GO:0070637	pyridine nucleoside metabolic process	2	1	0.002257
GO:0070638	pyridine nucleoside catabolic process	2	1	0.002257
GO:0042451	purine nucleoside biosynthetic process	64	2	0.002459
GO:0046129	purine ribonucleoside biosynthetic process	64	2	0.002459
GO:0019520	aldonic acid metabolic process	2	1	0.002573
GO:0006148	inosine catabolic process	3	1	0.003383
GO:0034356	NAD biosynthesis via nicotinamide riboside salvage pathway	3	1	0.003383

GO:0070970	interleukin-2 secretion	3	1	0.003422
GO:0042455	ribonucleoside biosynthetic process	83	2	0.004144
GO:0009163	nucleoside biosynthetic process	85	2	0.004334
GO:0034418	urate biosynthetic process	4	1	0.004508
GO:1901659	glycosyl compound biosynthetic process	87	2	0.004528
GO:0055086	nucleobase-containing small molecule metabolic process	334	3	0.006607
GO:0046102	inosine metabolic process	6	1	0.006752
GO:0019509	L-methionine biosynthetic process from methylthioadenosine	7	1	0.00787
GO:0043102	amino acid salvage	7	1	0.00787
GO:0071265	L-methionine biosynthetic process	7	1	0.00787
GO:0071267	L-methionine salvage	7	1	0.00787
GO:0072526	pyridine-containing compound catabolic process	7	1	0.007872
GO:0006166	purine ribonucleoside salvage	8	1	0.009072
GO:1901566	organonitrogen compound biosynthetic process	384	3	0.009443
GO:0050663	cytokine secretion	8	1	0.009787
GO:0046637	regulation of alpha-beta T cell differentiation	9	1	0.010834
GO:0046638	positive regulation of alpha-beta T cell differentiation	9	1	0.010834
GO:0019835	cytolysis	10	1	0.011225
GO:0009435	NAD biosynthetic process	10	1	0.011506
GO:0046634	regulation of alpha-beta T cell activation	10	1	0.012196
GO:0046635	positive regulation of alpha-beta T cell activation	10	1	0.012196
GO:0042102	positive regulation of T cell proliferation	11	1	0.012416
GO:0045582	positive regulation of T cell differentiation	11	1	0.013405

GO:0032946	positive regulation of mononuclear cell proliferation	12	1	0.013587
GO:0050671	positive regulation of lymphocyte proliferation	12	1	0.013587
GO:0070665	positive regulation of leukocyte proliferation	12	1	0.013587
GO:0019359	nicotinamide nucleotide biosynthetic process	12	1	0.014249
GO:0046130	purine ribonucleoside catabolic process	13	1	0.014696
GO:0019363	pyridine nucleotide biosynthetic process	13	1	0.015364
GO:0006152	purine nucleoside catabolic process	14	1	0.015809
GO:0006163	purine nucleotide metabolic process	167	2	0.016299
GO:0045621	positive regulation of lymphocyte differentiation	14	1	0.016948
GO:0006555	methionine metabolic process	15	1	0.016995
GO:0009086	methionine biosynthetic process	15	1	0.016995
GO:0045580	regulation of T cell differentiation	14	1	0.017002
GO:0009165	nucleotide biosynthetic process	174	2	0.01738
GO:0019674	NAD metabolic process	15	1	0.017677
GO:1901293	nucleoside phosphate biosynthetic process	177	2	0.017926
GO:0006195	purine nucleotide catabolic process	16	1	0.018098
GO:0046415	urate metabolic process	16	1	0.018199
GO:0042129	regulation of T cell proliferation	16	1	0.018278
GO:0044275	cellular carbohydrate catabolic process	15	1	0.018456
GO:0044711	single-organism biosynthetic process	903	4	0.018647
GO:0006716	juvenile hormone metabolic process	17	1	0.019012
GO:0006718	juvenile hormone biosynthetic process	17	1	0.019012
GO:0043174	nucleoside salvage	17	1	0.019154
GO:0050870	positive regulation of T cell activation	16	1	0.019278
GO:1903039	positive regulation of leukocyte cell-cell adhesion	16	1	0.019278

GO:0072525	pyridine-containing compound biosynthetic process	17	1	0.019857
GO:0042742	defense response to bacterium	184	2	0.020102
GO:0034112	positive regulation of homotypic cell-cell adhesion	17	1	0.020385
GO:0015771	trehalose transport	18	1	0.0205
GO:0045619	regulation of lymphocyte differentiation	17	1	0.020534
GO:0042454	ribonucleoside catabolic process	19	1	0.021353
GO:0032944	regulation of mononuclear cell proliferation	19	1	0.021656
GO:0050670	regulation of lymphocyte proliferation	19	1	0.021656
GO:0070663	regulation of leukocyte proliferation	19	1	0.021656
GO:0016106	sesquiterpenoid biosynthetic process	19	1	0.022195
GO:0015766	disaccharide transport	20	1	0.022942
GO:0015772	oligosaccharide transport	20	1	0.022942
GO:0009617	response to bacterium	200	2	0.023426
GO:0000097	sulfur amino acid biosynthetic process	22	1	0.024828
GO:0006714	sesquiterpenoid metabolic process	22	1	0.025504
GO:0009164	nucleoside catabolic process	24	1	0.027164
GO:0015985	energy coupled proton transport, down electrochemical gradient	24	1	0.02729
GO:0015986	ATP synthesis coupled proton transport	24	1	0.02729
GO:1901135	carbohydrate derivative metabolic process	572	3	0.02764
GO:0009067	aspartate family amino acid biosynthetic process	25	1	0.028103
GO:0051251	positive regulation of lymphocyte activation	24	1	0.028731
GO:0016114	terpenoid biosynthetic process	25	1	0.028902
GO:0002696	positive regulation of leukocyte activation	25	1	0.029829
GO:0050867	positive regulation of cell activation	25	1	0.029829
GO:0050863	regulation of T cell activation	25	1	0.029909

GO:1903037	regulation of leukocyte cell-cell adhesion	25	1	0.029909
GO:0006754	ATP biosynthetic process	29	1	0.032773
GO:1902107	positive regulation of leukocyte differentiation	27	1	0.032809
GO:0046496	nicotinamide nucleotide metabolic process	28	1	0.033792
GO:0098542	defense response to other organism	241	2	0.03401
GO:0009166	nucleotide catabolic process	30	1	0.034723
GO:0019362	pyridine nucleotide metabolic process	29	1	0.034887
GO:1901658	glycosyl compound catabolic process	31	1	0.035076
GO:0000096	sulfur amino acid metabolic process	32	1	0.035823
GO:1901292	nucleoside phosphate catabolic process	32	1	0.036904
GO:0009206	purine ribonucleoside triphosphate biosynthetic process	34	1	0.038266
GO:0006082	organic acid metabolic process	653	3	0.039178
GO:0009145	purine nucleoside triphosphate biosynthetic process	35	1	0.039361
GO:0009117	nucleotide metabolic process	268	2	0.039425
GO:0044283	small molecule biosynthetic process	273	2	0.039693
GO:0006753	nucleoside phosphate metabolic process	271	2	0.040191
GO:0009201	ribonucleoside triphosphate biosynthetic process	36	1	0.040788
GO:0006334	nucleosome assembly	31	1	0.040837
GO:0051249	regulation of lymphocyte activation	35	1	0.041597
GO:0006144	purine nucleobase metabolic process	36	1	0.042703
GO:0009306	protein secretion	36	1	0.043665
GO:1901137	carbohydrate derivative biosynthetic process	287	2	0.044164
GO:0008299	isoprenoid biosynthetic process	39	1	0.044197
GO:0072524	pyridine-containing compound metabolic process	38	1	0.044715

GO:0072523	purine-containing compound catabolic process	40	1	0.046283
GO:0022409	positive regulation of cell-cell adhesion	39	1	0.046597
GO:1902105	regulation of leukocyte differentiation	39	1	0.047442
GO:0009066	aspartate family amino acid metabolic process	43	1	0.047834
GO:0034110	regulation of homotypic cell-cell adhesion	40	1	0.047843
GO:0006733	oxidoreduction coenzyme metabolic process	42	1	0.049018
GO:0009142	nucleoside triphosphate biosynthetic process	44	1	0.049459

Table 3b. GO term up regulated in the category Cellular component

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0000786	nucleosome	17	2	0.000192
GO:1990104	DNA bending complex	17	2	0.000192
GO:0044815	DNA packaging complex	22	2	0.000352
GO:0032993	protein-DNA complex	33	2	0.0007
GO:0045261	proton-transporting ATP synthase complex, catalytic core F(1)	12	1	0.01361
GO:0033178	proton-transporting two-sector ATPase complex, catalytic domain	33	1	0.037039

Table3c. GO term up regulated in Gene Ontology category Molecular functions

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0016763	transferase activity, transferring pentosyl groups	27	2	0.000447
GO:0017061	S-methyl-5-thioadenosine phosphorylase activity	3	1	0.00338
GO:0002060	purine nucleobase binding	3	1	0.003383
GO:0004731	purine-nucleoside phosphorylase activity	3	1	0.003383
GO:0042301	phosphate ion binding	4	1	0.004507
GO:0002054	nucleobase binding	4	1	0.004648
GO:0008422	beta-glucosidase activity	9	1	0.010515
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	147	2	0.012497
GO:0003796	lysozyme activity	12	1	0.013456
GO:0016798	hydrolase activity, acting on glycosyl bonds	162	2	0.015102
GO:0046933	proton-transporting ATP synthase activity, rotational mechanism	14	1	0.015848
GO:0047886	farnesol dehydrogenase activity	15	1	0.016793
GO:0015151	alpha-glucoside transmembrane transporter activity	18	1	0.0205
GO:0015574	trehalose transmembrane transporter activity	18	1	0.0205
GO:0042947	glucoside transmembrane transporter activity	18	1	0.0205
GO:0015154	disaccharide transmembrane transporter activity	20	1	0.022942
GO:0015157	oligosaccharide transmembrane transporter activity	20	1	0.022942
GO:0016757	transferase activity, transferring glycosyl groups	202	2	0.02329

GO:0004568	chitinase activity	29	1	0.032335
GO:0030145	manganese ion binding	30	1	0.034791
GO:0008144	drug binding	30	1	0.035744
GO:0044769	ATPase activity, coupled to transmembrane movement of ions, rotational mechanism	35	1	0.039486
GO:0015926	glucosidase activity	37	1	0.042641
GO:0051119	sugar transmembrane transporter activity	43	1	0.048434

Table 4. GO term down regulated in females mated with lure-fed males (BP, Biological processes; MF, Molecular functions; CC, Cellular component)

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0050789	regulation of biological process (BP)	5387	1	0.00232
GO:0050794	regulation of cellular process (BP)	5044	1	0.00422
GO:0065007	biological regulation (BP)	5805	2	0.007141
GO:0005515	protein binding (MF)	3223	0	0.009313
GO:0019222	regulation of metabolic process (BP)	3216	0	0.010203
GO:0044464	cell part (CC)	10245	8	0.014812
GO:0031323	regulation of cellular metabolic process (BP)	2804	0	0.020083
GO:0060255	regulation of macromolecule metabolic process (BP)	2773	0	0.020886
GO:0080090	regulation of primary metabolic process (BP)	2660	0	0.025073
GO:0032502	developmental process (BP)	3787	1	0.030784
GO:0005575	cellular_component (CC)	11673	11	0.036444
GO:0044424	intracellular part (CC)	8928	7	0.041381
GO:0044767	single-organism developmental process (BP)	3523	1	0.0436

Table 5a. GO terms enriched virgin vs lure-fed mated – Biological processes – condition A enriched

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0030260	entry into host cell	56	2	0.000692
GO:0044409	entry into host	56	2	0.000692
GO:0046718	viral entry into host cell	56	2	0.000692
GO:0051806	entry into cell of other organism involved in symbiotic interaction	56	2	0.000692
GO:0051828	entry into other organism involved in symbiotic interaction	56	2	0.000692
GO:0052126	movement in host environment	56	2	0.000692
GO:0052192	movement in environment of other organism involved in symbiotic interaction	56	2	0.000692
GO:0051701	interaction with host	72	2	0.001189
GO:0018143	nucleic acid-protein covalent cross-linking	4	1	0.002865
GO:0018144	RNA-protein covalent cross-linking	4	1	0.002865
GO:0035897	proteolysis in other organism	4	1	0.002865
GO:0035915	pore formation in membrane of other organism	4	1	0.002865
GO:0039537	suppression by virus of host viral-induced cytoplasmic pattern recognition receptor signaling pathway	4	1	0.002865
GO:0039545	suppression by virus of host MAVS activity	4	1	0.002865
GO:0039546	suppression by virus of host MAVS activity by MAVS proteolysis	4	1	0.002865
GO:0039707	pore formation by virus in membrane of host cell	4	1	0.002865
GO:0044657	pore formation in membrane of other organism during symbiotic interaction	4	1	0.002865
GO:0044658	pore formation in membrane of host by symbiont	4	1	0.002865
GO:0051673	membrane disruption in other organism	4	1	0.002865
GO:0052025	modification by symbiont of host cell membrane	4	1	0.002865
GO:0052043	modification by symbiont of host cellular component	4	1	0.002865
GO:0052111	modification by symbiont of host structure	4	1	0.002865
GO:0052185	modification of structure of other organism involved in symbiotic interaction	4	1	0.002865
GO:0052188	modification of cellular component in other organism involved in symbiotic interaction	4	1	0.002865
GO:0052332	modification by organism of membrane in other organism involved in symbiotic interaction	4	1	0.002865
GO:0020012	evasion or tolerance of host immune response	5	1	0.003646
GO:0030683	evasion or tolerance by virus of host immune response	5	1	0.003646

GO:0039503	suppression by virus of host innate immune response	5	1	0.003646
GO:0044414	suppression of host defenses	5	1	0.003646
GO:0044501	modulation of signal transduction in other organism	5	1	0.003646
GO:0051805	evasion or tolerance of immune response of other organism involved in symbiotic interaction	5	1	0.003646
GO:0051833	suppression of defenses of other organism involved in symbiotic interaction	5	1	0.003646
GO:0052027	modulation by symbiont of host signal transduction pathway	5	1	0.003646
GO:0052029	negative regulation by symbiont of host signal transduction pathway	5	1	0.003646
GO:0052037	negative regulation by symbiont of host defense response	5	1	0.003646
GO:0052167	modulation by symbiont of host innate immune response	5	1	0.003646
GO:0052170	negative regulation by symbiont of host innate immune response	5	1	0.003646
GO:0052250	modulation of signal transduction in other organism involved in symbiotic interaction	5	1	0.003646
GO:0052261	suppression of defense response of other organism involved in symbiotic interaction	5	1	0.003646
GO:0052306	modulation by organism of innate immune response in other organism involved in symbiotic interaction	5	1	0.003646
GO:0052309	negative regulation by organism of innate immune response in other organism involved in symbiotic interaction	5	1	0.003646
GO:0052493	negative regulation by organism of signal transduction in other organism involved in symbiotic interaction	5	1	0.003646
GO:0052552	modulation by organism of immune response of other organism involved in symbiotic interaction	5	1	0.003646
GO:0052553	modulation by symbiont of host immune response	5	1	0.003646
GO:0052561	negative regulation by organism of immune response of other organism involved in symbiotic interaction	5	1	0.003646
GO:0052562	negative regulation by symbiont of host immune response	5	1	0.003646
GO:0052564	response to immune response of other organism involved in symbiotic interaction	5	1	0.003646
GO:0052572	response to host immune response	5	1	0.003646
GO:0075109	modulation by symbiont of host receptor-mediated signal transduction	5	1	0.003646
GO:0075111	negative regulation by symbiont of host receptor-mediated signal transduction	5	1	0.003646
GO:0075528	modulation by virus of host immune response	5	1	0.003646
GO:0030682	evasion or tolerance of host defense	6	1	0.004394

	response			
GO:0051807	evasion or tolerance of defense response of other organism involved in symbiotic interaction	6	1	0.004394
GO:0052031	modulation by symbiont of host defense response	6	1	0.004426
GO:0052255	modulation by organism of defense response of other organism involved in symbiotic interaction	6	1	0.004426
GO:0039506	modulation by virus of host molecular function	6	1	0.004426
GO:0039507	suppression by virus of host molecular function	6	1	0.004426
GO:0052055	modulation by symbiont of host molecular function	6	1	0.004426
GO:0052056	negative regulation by symbiont of host molecular function	6	1	0.004426
GO:0019049	evasion or tolerance of host defenses by virus	6	1	0.004426
GO:0044413	avoidance of host defenses	7	1	0.005174
GO:0044415	evasion or tolerance of host defenses	7	1	0.005174
GO:0051832	avoidance of defenses of other organism involved in symbiotic interaction	7	1	0.005174
GO:0051834	evasion or tolerance of defenses of other organism involved in symbiotic interaction	7	1	0.005174
GO:0039531	regulation of viral-induced cytoplasmic pattern recognition receptor signaling pathway	7	1	0.005185
GO:0039532	negative regulation of viral-induced cytoplasmic pattern recognition receptor signaling pathway	7	1	0.005185
GO:0044359	modulation of molecular function in other organism	7	1	0.005205
GO:0044362	negative regulation of molecular function in other organism	7	1	0.005205
GO:0052204	negative regulation of molecular function in other organism involved in symbiotic interaction	7	1	0.005205
GO:0052205	modulation of molecular function in other organism involved in symbiotic interaction	7	1	0.005205
GO:0052173	response to defenses of other organism involved in symbiotic interaction	8	1	0.005953
GO:0052200	response to host defenses	8	1	0.005953
GO:0075136	response to host	8	1	0.005953
GO:0050690	regulation of defense response to virus by virus	8	1	0.005984
GO:0016032	viral process	173	2	0.006809
GO:0044403	symbiosis, encompassing mutualism through parasitism	174	2	0.006889
GO:0044764	multi-organism cellular process	177	2	0.007121
GO:0019062	virion attachment to host cell	11	1	0.008313
GO:0044650	adhesion of symbiont to host cell	11	1	0.008313

GO:0044419	interspecies interaction between organisms	192	2	0.008379
GO:0050687	negative regulation of defense response to virus	12	1	0.00896
GO:0039656	modulation by virus of host gene expression	16	1	0.011631
GO:0039657	suppression by virus of host gene expression	16	1	0.011631
GO:0044406	adhesion of symbiont to host	17	1	0.012716
GO:0039694	viral RNA genome replication	17	1	0.01282
GO:0039703	RNA replication	17	1	0.01282
GO:0044033	multi-organism metabolic process	20	1	0.015139
GO:0019054	modulation by virus of host process	24	1	0.017809
GO:0002832	negative regulation of response to biotic stimulus	26	1	0.019421
GO:0019079	viral genome replication	26	1	0.019452
GO:0044068	modulation by symbiont of host cellular process	27	1	0.020118
GO:0019048	modulation by virus of host morphology or physiology	28	1	0.020886
GO:0002698	negative regulation of immune effector process	29	1	0.021758
GO:0044003	modification by symbiont of host morphology or physiology	31	1	0.023188
GO:0045824	negative regulation of innate immune response	32	1	0.024133
GO:0075732	viral penetration into host nucleus	37	1	0.026001
GO:1902594	multi-organism nuclear import	37	1	0.026001
GO:0019043	establishment of viral latency	42	1	0.029409
GO:0075713	establishment of integrated proviral latency	42	1	0.029409
GO:0075733	intracellular transport of virus	42	1	0.029622
GO:1902583	multi-organism intracellular transport	42	1	0.029622
GO:0044766	multi-organism transport	43	1	0.030384
GO:0046794	transport of virus	43	1	0.030384
GO:1902579	multi-organism localization	43	1	0.030384
GO:0032197	transposition, RNA-mediated	44	1	0.030657
GO:0051817	modification of morphology or physiology of other organism involved in symbiotic interaction	42	1	0.031476
GO:0043901	negative regulation of multi-organism process	44	1	0.03279
GO:0035821	modification of morphology or physiology of other organism	50	1	0.037293
GO:0050777	negative regulation of immune response	54	1	0.040416
GO:0019076	viral release from host cell	60	1	0.042189
GO:0050688	regulation of defense response to virus	56	1	0.042213
GO:0034645	cellular macromolecule biosynthetic process	1948	4	0.046416
GO:0031348	negative regulation of defense response	64	1	0.047298

Table 5b. GO terms enriched virgin vs lure-fed mated – Cellular component

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0033646	host intracellular part	56	2	0.000723
GO:0033643	host cell part	65	2	0.000982
GO:0044217	other organism part	65	2	0.000982
GO:0039661	host organelle outer membrane	4	1	0.002865
GO:0044162	host cell cytoplasmic vesicle membrane	4	1	0.002865
GO:0044191	host cell mitochondrial membrane	4	1	0.002865
GO:0044193	host cell mitochondrial outer membrane	4	1	0.002865
GO:0044384	host outer membrane	4	1	0.002865
GO:0044385	integral to membrane of host cell	4	1	0.002865
GO:0033655	host cell cytoplasm part	11	1	0.008015
GO:0019028	viral capsid	17	1	0.012645
GO:0033644	host cell membrane	26	1	0.019099
GO:0044218	other organism cell membrane	26	1	0.019099
GO:0044279	other organism membrane	26	1	0.019099
GO:0033647	host intracellular organelle	41	1	0.029108
GO:0033648	host intracellular membrane-bounded organelle	41	1	0.029108
GO:0042025	host cell nucleus	41	1	0.029108
GO:0000943	retrotransposon nucleocapsid	42	1	0.029335
GO:0030430	host cell cytoplasm	43	1	0.030952
GO:0044423	virion part	54	1	0.039017

Table 5c. GO terms enriched virgin vs lure-fed mated – Molecular function

GO Category	GO Function	Numbers in category	Number differentially expressed	P value
GO:0042302	structural constituent of cuticle	59	5	2.98E-10
GO:0005198	structural molecule activity	520	7	1.40E-08
GO:0003968	RNA-directed RNA polymerase activity	20	1	0.015112
GO:0004197	cysteine-type endopeptidase activity	39	1	0.029297
GO:0034062	RNA polymerase activity	63	1	0.046502
GO:0003724	RNA helicase activity	63	1	0.047453

Table 6. GO terms down regulated virgin vs lure-fed mated – condition A depleted

GO Category	GO Function	Numbers in category	Number differentially expressed	P value	
Biological processes					
GO:0031323	regulation of cellular metabolic process	2804	1	0.000548	BP
GO:0060255	regulation of macromolecule metabolic process	2773	1	0.000557	BP
GO:0080090	regulation of primary metabolic process	2660	1	0.00085	BP
GO:0007165	signal transduction	1883	0	0.00187	BP
GO:0048519	negative regulation of biological process	2116	1	0.002111	BP
GO:0019222	regulation of metabolic process	3216	3	0.002131	BP
GO:0048523	negative regulation of cellular process	1911	1	0.004281	BP
GO:0010468	regulation of gene expression	2204	1	0.006921	BP
GO:0009653	anatomical structure morphogenesis	1407	0	0.008501	BP
GO:0051171	regulation of nitrogen compound metabolic process	2127	1	0.008567	BP
GO:0009889	regulation of biosynthetic process	2100	1	0.009541	BP
GO:0031326	regulation of cellular biosynthetic process	2089	1	0.009869	BP
GO:0051128	regulation of cellular component organization	1252	0	0.011303	BP
GO:0019219	regulation of nucleobase-containing compound metabolic process	2049	1	0.011459	BP
GO:0010556	regulation of macromolecule biosynthetic process	1990	1	0.012866	BP
GO:2000112	regulation of cellular macromolecule biosynthetic process	1969	1	0.013722	BP
GO:0051252	regulation of RNA metabolic process	1866	1	0.018578	BP
GO:0031325	positive regulation of cellular metabolic process	1263	0	0.021833	BP
GO:0009892	negative regulation of metabolic process	1126	0	0.021975	BP
GO:2001141	regulation of RNA biosynthetic process	1801	1	0.022084	BP
GO:0006355	regulation of transcription, DNA-templated	1792	1	0.022661	BP
GO:0010604	positive regulation of macromolecule metabolic process	1150	0	0.027814	BP
GO:0007166	cell surface receptor signaling pathway	1113	0	0.028326	BP
GO:0048522	positive regulation of cellular process	2084	2	0.028391	BP
GO:0006928	cellular component movement	962	0	0.031253	BP
GO:0010605	negative regulation of macromolecule metabolic process	989	0	0.037512	BP

GO:0051246	regulation of protein metabolic process	807	0	0.037531	BP
GO:0031324	negative regulation of cellular metabolic process	933	0	0.038388	BP
GO:0044267	cellular protein metabolic process	1818	2	0.040692	BP
GO:0019538	protein metabolic process	2370	3	0.041767	BP
GO:0032268	regulation of cellular protein metabolic process	756	0	0.044279	BP
GO:0032879	regulation of localization	858	0	0.0496	BP

Cellular component

GO:0016020	membrane	3753	4	0.006526	CC
GO:0043227	membrane-bounded organelle	5567	8	0.006829	CC
GO:0043231	intracellular membrane-bounded organelle	5261	8	0.016174	CC
GO:0044425	membrane part	3809	5	0.038189	CC
GO:0005886	plasma membrane	1655	1	0.048329	CC

Appendix III: Draft MS, Comparative fruit fly ecology

Title: No evidence of niche segregation in *Bactrocera neohumeralis* (Diptera: Tephritidae), a sibling to the Queensland fruit fly, *Bactrocera tryoni*

Nagalingam Kumaran[1], Ema Tora Vueti[1], Brendan Missenden[2], Lara Senior[2], Megan Higgie[3] and Anthony R Clarke[1, 4]

1. School of Earth, Environmental and Biological Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Qld, Australia
2. Department of Agriculture and Fisheries, Ecosciences Precinct, Dutton Park, Queensland, Australia
3. College of Marine & Environmental Sciences, James Cook University, Queensland 4811, Australia
4. Plant Biosecurity Cooperative Research Centre, LPO Box 5012, Bruce, ACT, Australia

*Correspondence

N. Kumaran

School of Earth, Environmental and Biological Sciences,

Queensland University of Technology,

GPO Box 2434, Brisbane, Qld, Australia

k.nagalingam@qut.edu.au

For Journal: Bulletin of Entomological Research

Word count:

No evidence of niche segregation in *Bactrocera neohumeralis* (Diptera: Tephritidae), a sibling to the Queensland fruit fly, *Bactrocera tryoni*

Nagalingam Kumaran, Ema Tora Vueti, Brendan Missenden, Lara Senior, Megan Higgie and Anthony R Clarke

Abstract

Coexisting species differ in their ecological requirements to avoid competitive exclusion. However, species may also compete for common resources, without being limited simply by its availability, because they differ in the way they utilize the resources. Australian sibling species Bactrocera neohumeralis and Bactrocera tryoni occur sympatrically with no limitations. We tested whether these species differ in their resource utilization. We tested seasonal abundance, host use and habitat use by these two species. The historical seasonal abundance data revealed no difference in seasonal abundance with similar trend of increase and decrease in both species. Both species followed similar seasonal variation in all locations with greater number flies trapped during Sep-Nov and Feb-March in most of the locations. There were greater number of flies trapped in Dry sclerophyll forests followed by Horticulture farming system, and this habitat use pattern was similar in *B. tryoni and B. neohumeralis*. Host use did not differ between *B. tryoni* and *B. neohumeralis* with similar pattern of increase and decrease in all host plants. The data suggested no evidence of niche segregation in these species and supported the notion that *B. tryoni* and *B. neohumeralis* co-occur sympatrically with similar ecological requirements and resource utilization.

Introduction

Coexisting species differ to some extent in their ecological requirements to avoid competitive exclusion (Rathcke 1976). On the other hand, species may compete for a common resource, without being limited simply by its availability, because they differ in the way they utilize the resources (Heithaus et al. 1975). Evidences also suggest that, although the maximum niche overlap (niche supporting two species without any harm) theory is untestable, the niche supports species until the availability of resources become limited to one species in a coexistence scenario (Paine et al. 1981). When resources start to limit, then speciation is likely. In sympatric speciation for instance, resource utilization by one species lead to speciation in a close related coexisting sibling species (Berlocher & Feder 2002). In insects such speciation is well-known which leads to adaptation to new hosts through host shift, for example in the apple maggot fly *Rhagoletis pomonella* (Bush 1969).

Australian tephritid species *Bactrocera tryoni* and *B. neohumeralis* present yet another alike-example for sympatric speciation. In fact these two species are siblings with a very little variation in the morphology and reproduction. *Bactrocera neohumeralis* has brown numeral calli which is bright yellow in B. tryoni, and the former copulate at bright light during the midday while latter mate at dusk when light intensity is low (Birch & Vogt 1970; Smith 1979). Apart from these subtle variations, there are no other differences noticed until now. Flies with intermediate colour with part yellow and part brown humeral calli were found in wild, which were first regarded as colour variants (Hardy, 1951), then were long thought to represent hybrids (Bateman 1958, Birch 1961, Gibbs, 1965, Lewontin & Birch 1966, Wolda 1967a, b; Gibbs 1968; Pike et al. 2003); however the most recent microsatellite analysis indicates that these "hybrids" are intraspecies variants (Pike & Meats 2002; Gilchrist & Ling 2006). A set of 30 microsatellites reveals differentiation between *B. tryoni* and *B. neohumeralis* at the level of allele frequency, but no fixed microsatellite allele differences have been identified [Wang et al 2003].

Sharing of polymorphisms in coding and non-coding regions is also common indicating continued genetic exchange between the two species. Further studies on neo also evidenced lesser variations based on mating compatibility and genetic structure. Despite their very close genetic similarity (Morrow *et al.* 2000), *B. tryoni* and *B.neohumeralis* have very different pest status. Their recorded host lists are similar (Hancock *et al.* 2000), but *B. tryoni* is the major pest fruit fly for all of eastern Australia, while *B.neohumeralis* is, at worst, a pest of the tropics and subtropics (Drew *et al.* 1978).

Despite the systematic studies on morphological characters and microsatellites, and observations on mating behaviour, there are limited studies, barely one (comparative) study on the host use, with no other studies on ecological parameters such as habitat use and seasonal abundance. These studies are important to understand the functional variations between these two siblings. Gibbs (Gibbs 1965; 1967) carried out comparative studies on the host use of the two species in Rockhampton and concluded that inter-species competition was not the answer, while Meats (2006) concluded that an inability to handle cold did not restrict the southern range of *B. neohumeralis*. No other direct comparative ecological studies have been carried out on the two species.

We compared trap data collected on species abundance for both species in different landscapes, their seasonal abundance over several years and abundance in different host fruits to investigate evidence of niche separation between these two sibling species. We summarize here the similarities in abundance of these two species at time scale, spatial scale and resource use. The data showed no evidence of difference between two species in terms of their habitat use, seasonal abundance and host use and provided no evidence of niche segregation. The data provide further support that both the species are

sympatric sibling species.

Materials and Methods

Seasonal abundance

Unpublished historical dataset from May's (1961) thesis was used for the comparison between B. tryoni and B. neohumeralis seasonal abundance. A detailed methodology used is described elsewhere (May 1961; Muthuthantri et al 2010). Briefly, 15 trapping sites across Queensland, Australia were used in this study, ranging from Stanthorpe in the south to Cairns in the north of Queensland, Australia (Table 1). The sites vary significantly in their climate across this range and are explained in the companion paper (Muthuthantri et al. 2010). It is important to keep in mind that, as the dataset are historical, the current usage of the land is different from what it was in the 1950s.

Detailed methodology on fly trapping is described in Muthuthantri et al (2010). Fly trapping was carried out for 2–7 years for any locations using a liquid lure in McPhail traps after being diluted with water (1:30 lure: water ratio) as described by May and Caldwell (1944). This lure formulation contained pulped orange (10 ounces (=283 g)), ammonium carbonate (0.5 ounces (=14.2 g)) and rain water (1 pint (=0.57L)), held for 24 h in a closed container. Totally Ten traps were maintained at each trapping station and all trapping stations were located in fruit growing areas where traps were placed in host trees. Traps were hung permanently in host trees and serviced weekly and trapped flies were identified at species level. For this paper, data on B. tryoni and B. neohumeralis only were extracted and used.

Where raw data are available (Atherton, Ayr, Rockhampton and Lawes), descriptive analyses were redone. For trap locations Maryborough, Sunny bank, Kamerunga and Gayndah, mean percent of total trap catches were only available and used as such for comparison. For South Johnston, mean seasonal data only available for winter, summer, autumn and spring which were used as such. For stations Rita Island, Nambour, Withcott, Toowoomba, Offham and Stanthorpe, only total trap catches data were available which is presented in table 1.

The trap data presented in May's (1961) thesis is stacked vertical columns for each week and for immature males, mature males, immature females, non-gravid females and mature females as a single column. This column data were transcribed from the graphs into Excel format and total number of flies by summing up all age class, sex and mating status were used to compare B. tryoni and B. neohumeralis. For locations, Atherton, Ayr, Rockhampton and Lawes, the weekly data were summed to a calendar month and presented as monthly data. While combining weekly data could hide potential weekly variation within species, we believe it is trivial since it is a comparative study between two species. The reader is referred to May's (1961) thesis, held by The University of Queensland library, if they wish to view weekly phenology data.

Abundance in different landscapes (Habitat use)

The study was undertaken to determine the distribution of B. tryoni and B. neohumeralis across various habitats. A total of six habitat types which included; grassland, residential, intensive horticultural farming, mixed sugarcane and other crops farming, dry forest and wet forest were chosen for the study in the Bundaberg region of Queensland (Habitats identified for trapping sites were chosen from human defined perspective and from previous research on fruit fly distributions in its natural environment (Drew and Hooper 1983; Raghu 1998; Raghu and Clarke 2001). Specific descriptions of each habitat type are outlined in Table S1. For each habitat, there were three replicates and three fruit fly trapping sites were

established in each replicate. A total of 54 trap sites were set up on a monthly basis.

Trapping commenced in September 2010 (spring) and completed in March 2011 using cuelure trap (modified Steiner). The trapping program involved setting up the traps on a monthly basis at each trap site, removing the traps at each site three days after setting up and replacing the traps with fresh wicks and attractants in the following month. Flies collected over the three days in each month was counted and identified to species level. Identifications of flies were carried out using the taxonomic keys in Memoirs of the Queensland Museum (Drew 1989). Fruit fly specimens that required identity confirmation were made by Prof R.A.I. Drew (Griffith University, Brisbane). To maintain ease of access and safety during the trapping program, some traps in the wet riverine forests were placed alongside the road. Where possible, most traps were hung on host fruit trees while some were hung on non-host trees, at least one and a half to two meters above the ground. A few traps were hung on the tallest shrub (in grasslands) or trellis or stick in row cropping (tomato, pumpkin, watermelon and capsicum) farms.

Host use pattern

Host use data were extracted from trapping programs held by [Queensland] Department of Agriculture, Brisbane, Australia. Fruit was collected from the DAF Maroochy Research Facility, Nambour, Queensland from separate orchard blocks that consisted of either mixed topical fruit varieties, Stone fruit or mangoes, as well as some isolated trees. Small numbers of ripe or mature green fruit were sampled on a weekly or fortnightly basis between December 2012 and October 2015. Fruit were selected from a number of trees within the orchard block and picked from various heights and aspects to obtain a random sample. Fruit that were of poor quality were not rejected from collections in order to prevent sampling bias.

Fruit were placed in paper bags and transported to the DAF laboratories. Fruit were counted and weighed and then placed on gauzed plastic containers over vermiculite, in plastic boxes with gauzed lids to allow ventilation. Boxes were held in a Controlled Environment Room (26°C and 70%RH) to allow insects to develop through to the pupal stage. The vermiculite was sieved weekly until all insects had exited fruit and pupated, and then fruit was inspected before being discarded. Fruit fly pupae that were reared out of fruit samples were placed into small plastic boxes with gauzed lids containing vermiculite. Once adult fruit flies had emerged and died they were identified to species level.

Results

Seasonal abundance

Bactrocera neohumeralis was found in all seasons as *B. tryoni* and followed similar trend in terms of abundance. While there was a difference in abundance with more number of B. tryoni trapped in almost all locations, there was a significant correlation in number of flies detected in most of the locations (Table 2). *B. neohumeralis* showed decreasing trend in abundance from North of Queensland to Southern part (Table 1).

The seasonal abundance of both species followed similar pattern in trap catches collected from Atherton, Ayr, Rockhampton, Kamerunga, Maryborough, Sunny bank, Gayndah, South Johnston and Lawes (Fig. 1 & 2). Mean number of flies trapped per year in various seasons at different locations are presented in table 1.

Flies were more abundant during October to December both in B. tryoni and B. neohumeralis at Atherton, Ayr, Kamerunga and Lawes. In Atherton, number of flies trapped in B. tryoni ranged from 55 to 631, and in *B. neohumeralis* in ranged between 2 and 119 (Fig. 1a). In Ayr, number of flies trapped in B. tryoni ranged from 23 to 389, and in *B. neohumeralis* it ranged between 8 and 110 (Fig. 1b) and the trend was similar for both *B. neohumeralis* and *B. tryoni*. In Lawes, number of B. tryoni and B. neohumeralis trapped ranged from 97 to 1257 and 0 to 31 respectively. In addition to October-December, more number of flies detected during February-April in Lawes. In Kamerunga, number of B. tryoni and B. neohumeralis trapped ranged from 2 to 422 and 4 to 86 respectively. In all locations, although B. tryoni is more abundant compared to B. neohumeralis, the trend was similar with positive correlation in seasonal abundance.

In Rackhampton, abundance of *B. tryoni* and *B. neohumeralis* ranged from 27 to 811 and 4 to 128 respectively. The abundance was peak during September both in *B. tryoni* and *B. neohumeralis*. In Maryborough and Sunny bank, flies were more abundant during September-October and January with low numbers during November-December.

In Gayndah, flies were abundant during September-November and February-March with low numbers detected during January. During December, B. tryoni was low, whereas B. neohumeralis is high. In south Johnston, more number of flies trapped during April - June both in B. tryoni and B. neohumeralis with numbers decreasing in July-September, October-December and January-March. B. tryoni and B. neohumeralis followed similar trend with positive correlation in abundance season-wise.

Habitat use

There was a significant difference in abundance of flies with more number of flies detected in Dry forest followed by Horticulture system. Abundance of flies in Dry forest significantly differed from Grassland, Wet forest, Mixed farming and Suburbia (B. tryoni: $F = 8.57$, $p < 0.001$; B. neohumeralis: $F = 9.65$, $p < 0.001$). There was no difference among Grassland, Mixed farming, Dry forest and Wet forest (Fig. 3).

There was a significant correlation between B. tryoni and B. neohumeralis in Grassland ($r = 0.841$, $p < 0.05$) and Mixed farming ($r = 0.990$, $p < 0.001$) habitats, but it was not significant for other habitats (Dry forest: $r = 0.063$, $p = 0.894$; Wet forest: $r = 0.351$, $p = 0.440$; Horticulture: $r = 0.276$, $p = 0.549$; Suburbia: $r = 0.335$, $p = 0.462$). Abundance in different landscapes showed similar trend for both B. tryoni and B. neohumeralis (Fig. 4).

Host use

Similar host use pattern by B. tryoni and B. neohumeralis noticed in white sapote, mulberry, peach, plum, carambola, nectarine, Feijoa and guava (Fig. 5). There was a significant correlation in host use by B. tryoni and B. neohumeralis in these host plants (Table). In guava, only two collections were made in September (1 B. tryoni and no B. neohumeralis) and October 2013 (1609 B. tryoni and 1353 B. neohumeralis). In hosts viz., sapodilla, syzigium, Grumichama, Hog plum, jabotica, white mulberry, avocado, black sapote, longan, mango and cashew, no B. heohumeralis was recovered. However, monthly mean number of B. tryoni recovered from most of these hosts itself is very low compared to other hosts (Sapodilla: 46.67±17.34; Grumichama: 105.67±92.81; Jabotica: 4.50±2.50; white mulberry: 6.00±0.00; Syzigium: 263.00±98.96; mango: 211.20±61.50; avocado: 68.00±0.00; black sapote: 47.00±21.00; hog plum: 183.00±105.52; cashew: 29.00±0.00; Longan: 1.00±0.00). In Feijoa, equal number (71 flies) of B. tryoni and B. neohumeralis was recovered only once throughout the collection

period. Although no neohumeralis was recovered from the above minor hosts, there was no evidence of presence of neohumeralis in hosts where tryoni is absent.

There was no significant difference among host preference both in B. tryoni ($F = 2.09$, $p = 0.083$) and B. neohumeralis ($F = 1.33$, $p = 0.268$). However, the trend in number of flies detected in different host fruits was similar for B. tryoni and B. neohumeralis with more number of flies detected from peach (Fig. 6).

Discussion

Summary of results

Abundance of *B. neohumeralis* and *B. tryoni* was highly correlated and followed similar trend. Both species followed similar seasonal variation in all locations with greater number flies trapped during Sep-Nov and Feb-March in most of the locations. With respect to the habitat use, there were greater number of flies trapped in Dry sclerophyll forests followed by Horticulture farming system, and this habitat use pattern was similar in B. tryoni and B. neohumeralis. Host use did not differ between B. tryoni and B. neohumeralis with similar pattern of increase and decrease in all host plants. However, there are few minor host plants, where only B. tryoni was observed to emerge not B. neohumeralis. Overall, these results showed no evidence that neohumeralis has a unique niche (seasons, habitats and host plants) where tryoni is absent or vice versa.

The results on host use are partly similar to the study by Gibbs (1967) who found that *B. tryoni* and B*. neohumeralis* preferred the same kind of fruits. Further, the preference observed was associated with the relative abundance of the species rather than strict host preference. For instance, at Rockhampton, where *B. tryoni* constituted 90% of population, it showed no preference for fruits, whereas B. neohumeralis preferred a few over others. At Yeppoon, where B*. neohumeralis* is prevalent, it did not show host preference but *B. tryoni* did. In addition, Gibbs (1967) found no evidence of competitive displacement between *B. tryoni* and *B. neohumeralis*, and evidenced both species preferred same kinds host fruits when a choice was available. It was clear further from his study that density of one species does not affect density of other.

Perhaps fruit preference in Gibb's study is because of the density of one species not the other. For instance, when density is low, host availability is high so flies execute preferences. Whereas when density of flies is high, hosts available should become less and do not execute any preference. This explains why B. neohumeralis is absent in minor hosts in the current study. It is perhaps because of low population density of B. neohumeralis, not a true preference for hosts.

Variation in abundance in Northern and Southern Queensland

Perhaps it is an effect of climate and species physiology. Birch (1961) found that when the two species from Cairns were reared together in crowded cages, D. *neohumeralis* was superior to D. *tvyoni* at 25°C but D. *tryoni* was superior at 22°C. This result supports the hypothesis that D. *neohumevalis* is better adapted to tropical climates. Bateman (1958) measured the ability of each species to withstand extreme temperatures. He found that D. *tryoni* adults survived significantly better than *D, neohumevalis* at -5, 0, and 40°C.

Evolution of mating system (tryoni is more evolved?)

It is likely that B. tryoni has been more quickly evolving than neohumeralis and been adapted to cooler

climatic conditions and selected to mate at dusk. Flies that mate at dusk perhaps selected the trait to avoid predation. The short mating window in tryoni and irregular, broad window in neo support this notion. If day mating is a trait selected by neo, it is unlikely to see any benefits of such selection. Probably day mating is not a behaviour that potentially maximise the reproduction. It is possible that they suffer far more predation than dusk maters like tryoni.

Mechanisms mediating mating isolation/speciation

It is not competitive displacement that is making tryoni more abundant, there may be other factors driving this abundance of one species over another (Andrewartha & Birch 1954). For instance, tryoni is more prevalent towards southern part of Queensland; this is perhaps because light intensity in the south may not be suitable for neohumeralis as it decreases from the equator.

Neohumeralis is abundant probably in places where wild hosts are predominant. Tryoni is adapted or made use of horticulture and expanded its range, whereas neo is not as adaptive as tryoni to new hosts and hence they are less prevalent in commercial hosts.

Gibbs (1967): Historical records of fruit flies in Queensland have been compiled by May (1963) and can provide a useful background to this study. In 1909, Froggatt reported varieties of D. *tvyoni* from tomatoes and other fruit in Queensland. May considers 1138 G. W. GIBBS these were probably D. neohumeralis. Other specimens collected by Tryon in cultivated and wild hosts prior to 1927 included both D. tryoni and D. neohumeralis. These records indicate that the present coexistence of the two species is not a temporary state, but has very likely existed for at least 40 years.

Displacement by B. tryoni: a medfly scenario

The coexistence of D. neohumeralis and ***D.*** tryoni in Queensland, is in contrast to the displacement of *Ceratitis capitata* by *D. tryoni* in New South Wales (Andrewartha and Birch 1954). In the latter area, *C. capitata*, an introduced species, was a serious pest of fruit until about 1940. *D. tryoni*, which now lives in almost indentical circumstances which were previously characteristic of *C. capitata*, has completely replaced it. Andrewartha and Birch stress that food and other requirements for the flies were not in short supply during the replacement. No explanation for the complete replacement has been discovered. D. neohumeralis is obviously better adapted to live with **D,** tryoni than *C. capitata* was, but the reasons for this are not known.

Character displacement

Character displacement explains why neo is not detected in some places where only tryoni detected. The theory predicts that differences among similar species will be more prominent in places where the species coexist; but the differences will be minimized or lost in places where the species distributions do not overlap. If this is the case, there will be clear difference (morphological and reproductive behavioural) between tryoni and neo in areas such as north of Queensland where both species coexist.

Acknowledgement

References

Andrewartha, H.G. and Birch, L.C. (1954). *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago. 782pp

Berlocher, S.H. & Feder, J.L. (2002). Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology* 47, 773-815.

Birch, L.C. (1961). Natural selection between two species of tephritid fruit fly of the genus *Dacus*. *Evolution*, 360-374.

Birch, L.C. & Vogt, W.G. (1970). Plasticity of taxonomic characters of the Queensland fruit flies *Dacus tryoni* and *Dacus neohumeralis* (Tephritidae).*Evolution* 24, 320-343.

Bush, G.L. (1969). Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution*, 237-251.

Gibbs, G.W. (1967). The comparative ecology of two closely related, sympatric species of *Dacus* (Diptera) in Queensland. *Australian Journal of Zoology* 15, 1123-1139.

Gilchrist, A.S. & Ling, A.E. (2006). DNA microsatellite analysis of naturally occurring colour intermediates between *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy) (Diptera: Tephritidae). *Australian Journal of Entomology* 45, 157-162.

Heithaus, E.R., Fleming, T.H. & Opler, P.A. (1975). Foraging patterns and resource utilization in seven species of bats in a seasonal tropical forest. *Ecology 56*, 841-854.

Lewontin, R.C. & Birch, L.C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution*, 315-336.

May, A.W.S. (1961)—A taxonomic and ecological study of the Dacinae (fam. Trypetidae) in Queensland. Thesis (Ph.D.) University of Queensland.

May, A. W. S., & Caldwell, N. E. H. (1944). Fruit fly control. *Queensland Agricultural Journal*, *58*, 224-229.

Meats, A. (2006). Attributes pertinent to overwintering potential do not explain why Bactrocera neohumeralis (Hardy)(Diptera: Tephritidae) does not spread further south within the geographical range of B. tryoni (Froggatt). *Australian Journal of Entomology*, *45*(1), 20-25.

Morrow, J., Scott, L., Congdon, B., Yeates, D., Frommer, M., & Sved, J. (2000). Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution*, *54*(3), 899-910.

Muthuthantri, S., Maelzer, D., Zalucki, M. P., & Clarke, A. R. (2010). The seasonal phenology of Bactrocera tryoni (Froggatt)(Diptera: Tephritidae) in Queensland. *Australian Journal of Entomology*, *49*(3), 221-233.

Paine, T. D., Birch, M. C., & Švihra, P. (1981). Niche breadth and resource partitioning by four sympatric species of bark beetles (Coleoptera: Scolytidae). *Oecologia*, *48*(1), 1-6.

Pike, N., & Meats, A. (2002). Potential for mating between Bactrocera tryoni (Froggatt) and Bactrocera neohumeralis (Hardy)(Diptera: Tephritidae).*Australian Journal of Entomology*, *41*(1), 70-74.

Pike, N., Wang, W. Y. S., & Meats, A. (2003). The likely fate of hybrids of Bactrocera tryoni and Bactrocera neohumeralis. *Heredity*, *90*(5), 365-370.

Rathcke, B. J. (1976). Competition and coexistence with a guild of herbivorous insects. *Ecology*, *57*(1), 76-87.

SMITH, P. H. (1979). Genetic manipulation of the circadian clock's timing of sexual behaviour in the Queensland fruit flies, Dacus tryoni and Dacus neohumeralis. *Physiological Entomology*, *4*(1), 71-78.

Wang, Y., Yu, H., Raphael, K., & Gilchrist, A. S. (2003). Genetic delineation of sibling species of the pest fruit fly Bactocera (Diptera: Tephritidae) using microsatellites. *Bulletin of entomological research*, *93*(04), 351-360.

Wolda, H. (1967). Reproductive isolation between two closely related species of the Queensland fruit fly Dracus tryoni (Frogg) and D. neohumeralis Hardy (Diptera: Tephritidae) II. Genetic variation in humeral callus pattern in each species as compared with laboratory-bred hybrids. *Australian Journal of Zoology*, *15*(3), 515-539.

Table 1 Prevalence of tryoni:neohumeralis

Stations	Years trap maintained	Tryoni per trap per year	Neo per trap per year	tryoni: neohumeralis
Kamerunga	4	48.6	14	3.47 : 1
Atherton	4	301.9	44.6	6.77 : 1
South Johnston	3	4.4	7.4	0.59 : 1
Rita island	3	37.6	25.1	1.5 : 1
Ayr	4	160.5	38.7	4.13 : 1
Rockhampton	3	222.1	36.8	6.04 : 1
Maryborough	3	34.97	7.9	4.43 : 1
Gayndah	2	550.1	6.6	83.35 : 1
Nambour	1	148.1	15.1	9.81 : 1
Sunnybank	3	734.1	59	25.31 : 1
Lawes	6	1065.3	14.2	75.02 : 1
Withcott	2	231.2	2.6	88.92 : 1
Toowoomba	6	313.98	3.34	94.01 : 1
Offham	1	39.2	0	-
Stanthorpe	4	191.7	0.125	1533.6 : 1

Table 2. Correlation on seasonal abundance of tryoni vs. neo at various trap stations in Queensland

Stations	**Pearson correlation ®**	***p***
Kamerunga	0.947	<0.001
Atherton	0.636	0.026
South Johnston	0.915	0.030
Ayr	0.946	<0.001
Rockhampton	0.959	<0.001
Maryborough	0.942	<0.001
Gayndah	0.612	0.034
Sunnybank	0.879	<0.001
Lawes	0.811	0.001

Table 3. Correlation

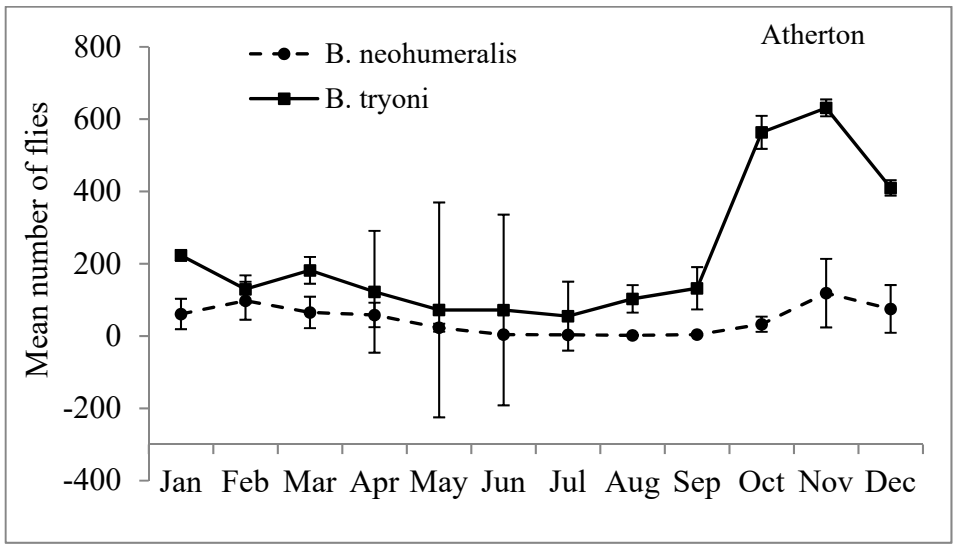
Host	Pearson correlation (r)	*p*
White sapote	0.979	<0.001
Mulberry	0.985	<0.001
Nectarine	0.950	<0.05
Plum	0.996	<0.001
Carambola	0.935	<0.001
Peach	0.992	<0.001
Guava	1.000	<0.001

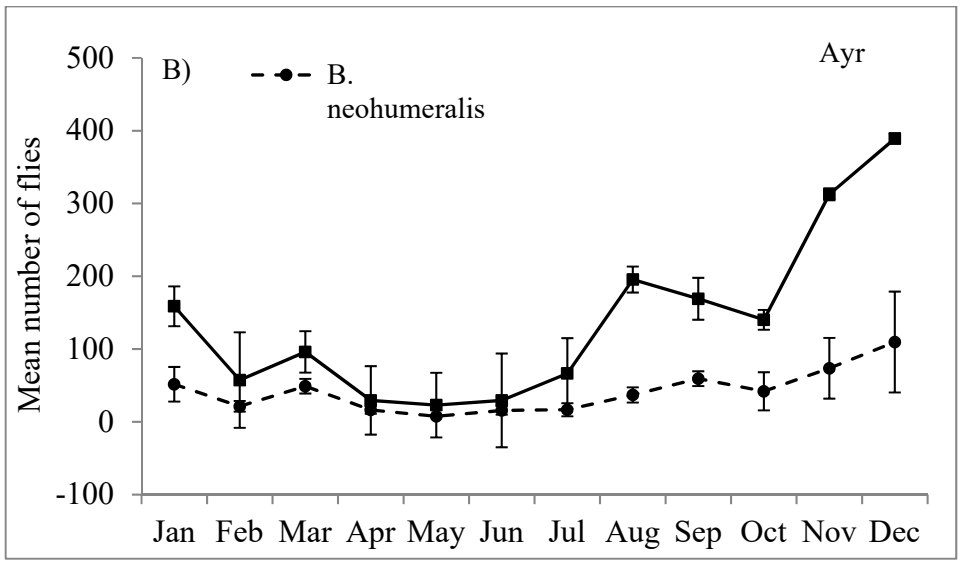
Table S1: Habitat descriptions of trap locations

Habitat type	Descriptions	Locations (latitude & longitude) T1 – trap site 1; T2-trap site 2; T3-trap site 3
Grasslands (Fig. 2.2)	This habitat included dry grazing and baron land and coastal wetland.	Price Street, West Bundaberg: *T1-S24°54'56.7" E152°18'23.1"T2-S24°55'03.4" E152°18'20.6"T3-S24°55'12.4" E152°18'17.3"* Bundaberg Port road, East Bundaberg: *T1-S24°46'15.3" E152°23'19.2"T2-S24°46'18.2" E152°23'22.5"T3-S24°46'19.9" E152°23'24.6"* Old Windermere road, S Bundaberg: *(T1-S24°53'20.6" E152°28'04.2"T2-S24°53'31.2" E152°28'05.6"T3-S24°53'43.3" E152°28'06.4"*
Suburbia/ Residential (Fig. 2.3)	Areas included residential suburbs in the coastal region, those less than two kilometres from the Burnett river and residential areas over two kilometres from the Burnett river.	Moore Park road, North Bundaberg (coastal): *T1-S24°43'07.5" E152°16'35.3"T2-S24°43'09.0" E152°16'48.1"T3- S24°43'25.4" E152°17'07.9"* West Bundaberg – Cox st., Diamond St& Powers st (near Burnett river): *T1-S24°87'17.3" E152°31'72.21"T2-S24°87'43.0" E152°31'60.0"T3- S24°87'79.3" E152°33'82.2"* East Bundaberg – Elliot Heads road (over 3km away from river and coastline) *T1-S24°88'08.3" E152°38'00.0"T2-S24°88'38.0" E152°38'30.7"T3- S24°88'52.3" E152°37'84.2"*
Intensive horticultural farming (Fig.2.4)	Farms include those that cultivated row crops such as tomatoes and zucchini and tropical fruit tree orchards such as mango, avocado, lychee, citrus, figs and bananas.	Moore Park road Farm 1 (north Bundaberg) *T1-S24°54'56.7" E152°18'23.1"T2-S24°55'03.4" E152°18'20.6"T3-S24°55'12.4" E152°18'17.3"* Moore Park road Farm 2 (north Bundaberg): *T1-S24°54'56.7" E152.18'23.1"T2-S24°55'03.4" E152°18'20.6"T3- S24°55'12.4" E152°18'17.3"* St. John Road Farm 3 (south Bundaberg): *T1-S24°52'00.9" E152°24'57.3"T2-S24°52'03.5" E152°24'56.7"T3-S24°52'06.0" E152°24'57.9"*
Mixed sugarcane & other crops farming (Fig. 2.5)	Farms selected were those that cultivated sugarcane as the main crop with other crops such as watermelons, pumpkins, sweet	Elliot Head's road Farm 1(East Bundaberg): *T1-S24°52'56.1" E152°23'36.0"T2-S24°52'43.7" E152°23'46.4"T3-S24°52'34.6" E152°23'47.8"* Elliot Head's road Farm 2 (South Bundaberg): *T1-S24°52'58.5" E152°27'01.1"T2-S24°53'06.9"*

	potatoes, strawberries, zucchini and tomatoes.	*E152°27'00.7"T3- S24°53'08.4" E152°26'46.4"* Mittelheusers rd, East Bundaberg: *T1-S24°47'08.3" E152°24'48.8"T2-S24°47'15.8" E152°24'46.5"T3-S24°47'21.1" E152°24'57.3"*
Dry sclerophyll forest (Fig. 2.6)	This habitat included sclerophyll forest over 2 kilometres away from the Burnett river. Forest canopy cover was between 40-55% (Walker and Hopkins 1990)	Bundaberg Ring road, west Bundaberg: *T1-S24°54'50.3" E152°20'19.3"T2-S24°54'42.4" E152°20'31.3"T3-S24°54'48.2" E152°20'48.2"* Norville park road, west Bundaberg: *T1-S24°53'25.5" E152°20'09.1"T2-S24°53'23.0" E152°20'19.9"T3-S24°53'22.4" E152°20'23.0"* Hummock road, south Bundaberg: *T1-S24°50'36.0" E152°25'33.6"T2-S24°50'33.2" E152°25'32.9"T3-S24°50'31.9" E152°25'35.2"*
Wet riverine sclerophyll forest (Fig.2.7)	The forests in this habitat were located less than one kilometre from the Burnett river or water courses. Vegetation consisted of the dense sclerophyll forest and eucalyptus species. Canopy cover is between 55-70% (Walker and Hopkins 1990)	Forest at the botanical gardens, north Bundaberg: *T1-S24°51'04.6" E152°20'15.3"T2-S24°51'10.8" E152°20'13.0"T3- S24°51'19.9" E152°20'11.4")* Queens Park, west Bundaberg: *T1-S24°52'02.7" E152°20'01.2"T2-S24°52'04.1" E152°19'54.1"T3-S24°52'04.7" E152°19'47.2")* Strathdees road, east Bundaberg: *T1-S24°47'35.7" E152°23'18.1"T2-S24°47'30.8" E152°23'19.4"T3-S24°47'16.6" E152°23'21.1"*

Figure 1A - C. Seasonal abundance





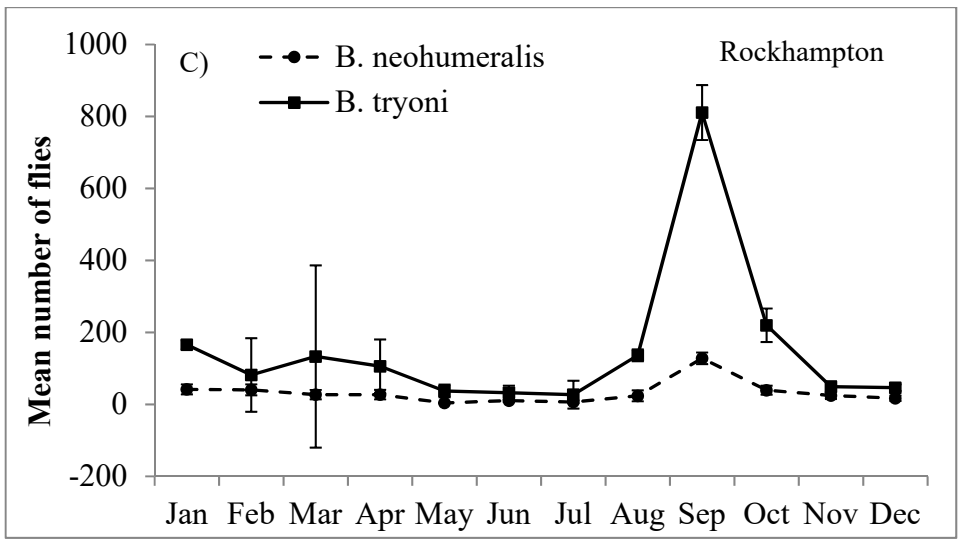


Figure 1D. Seasonal abundance

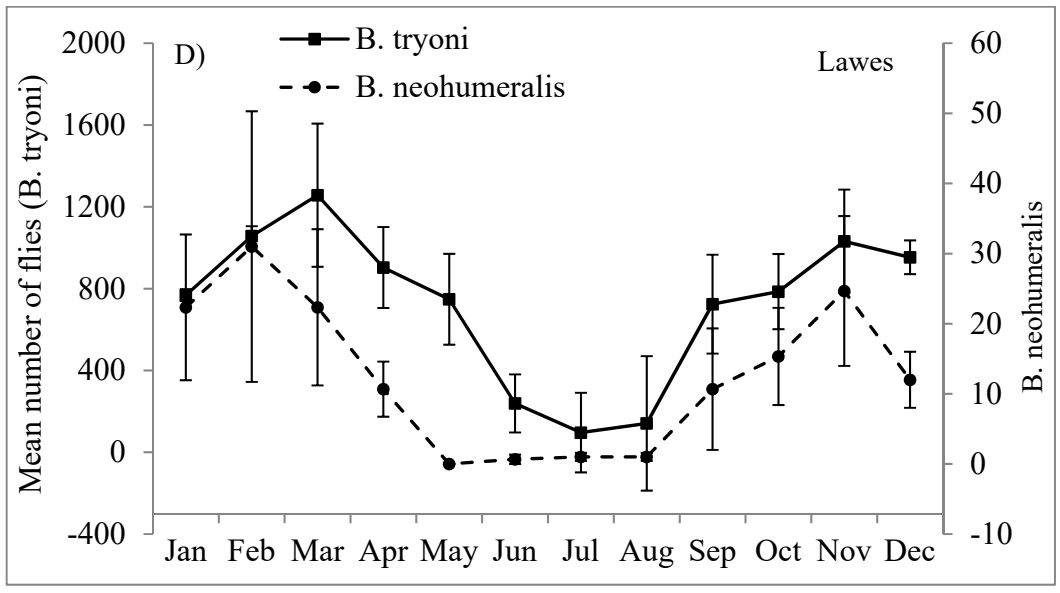
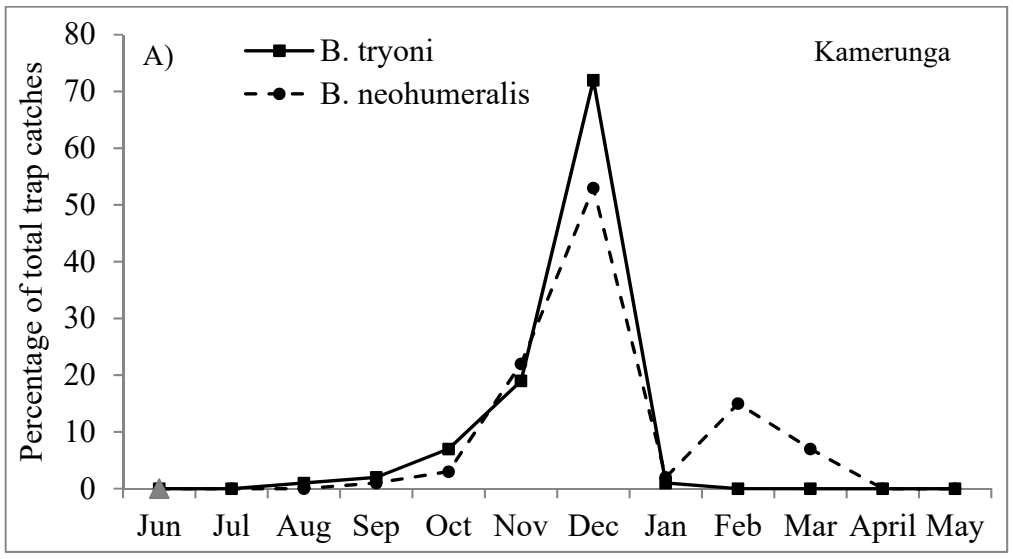
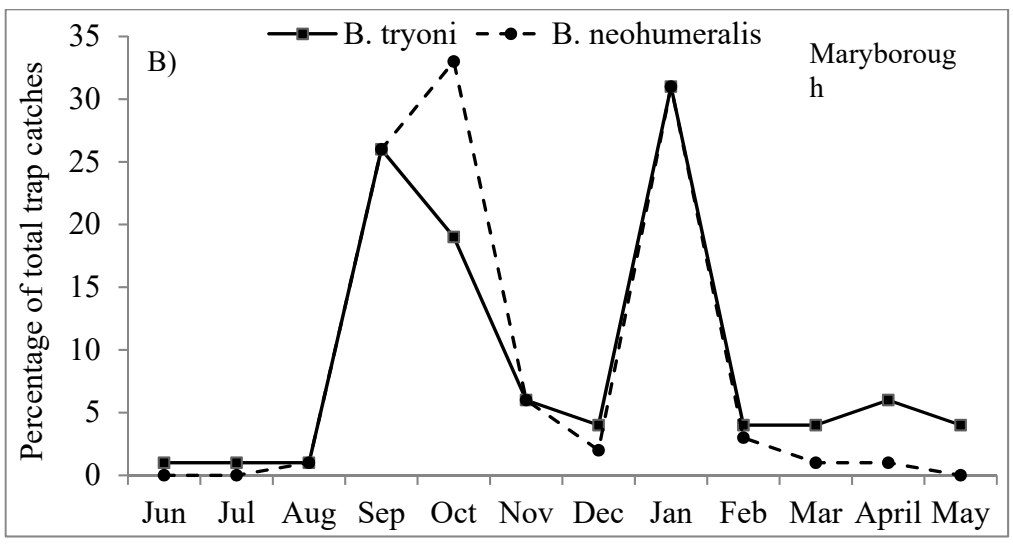


Figure 2A-C.


A)
B. tryoni
B. neohumeralis
Kamerunga
Percentage of total trap catches
80
70
60
50
40
30
20
10
0
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar April May


B)
B. tryoni
B. neohumeralis
Maryborough
Percentage of total trap catches
35
30
25
20
15
10
5
0
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar April May

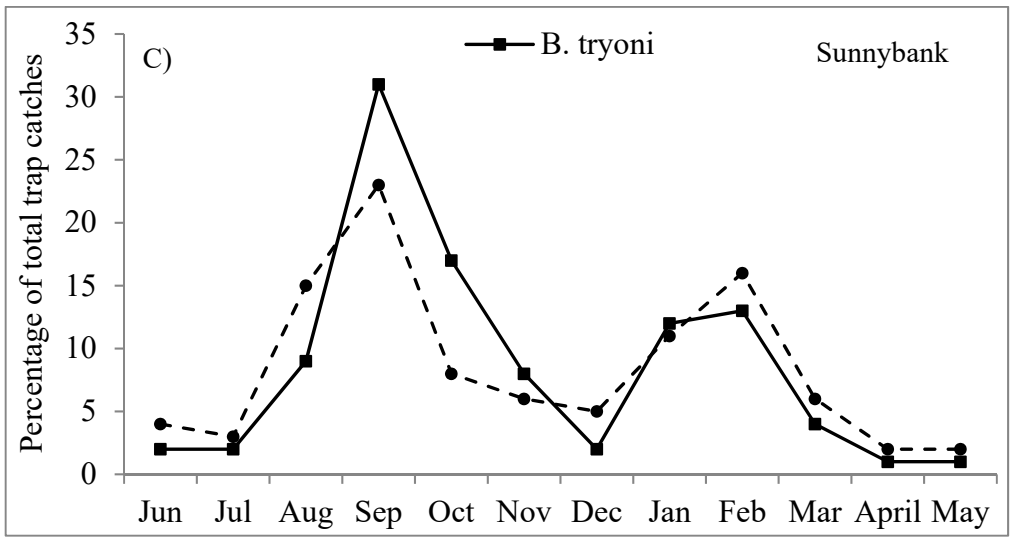

C)
B. tryoni
Sunnybank
Percentage of total trap catches
35
30
25
20
15
10
5
0
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar April May

Figure 2D.

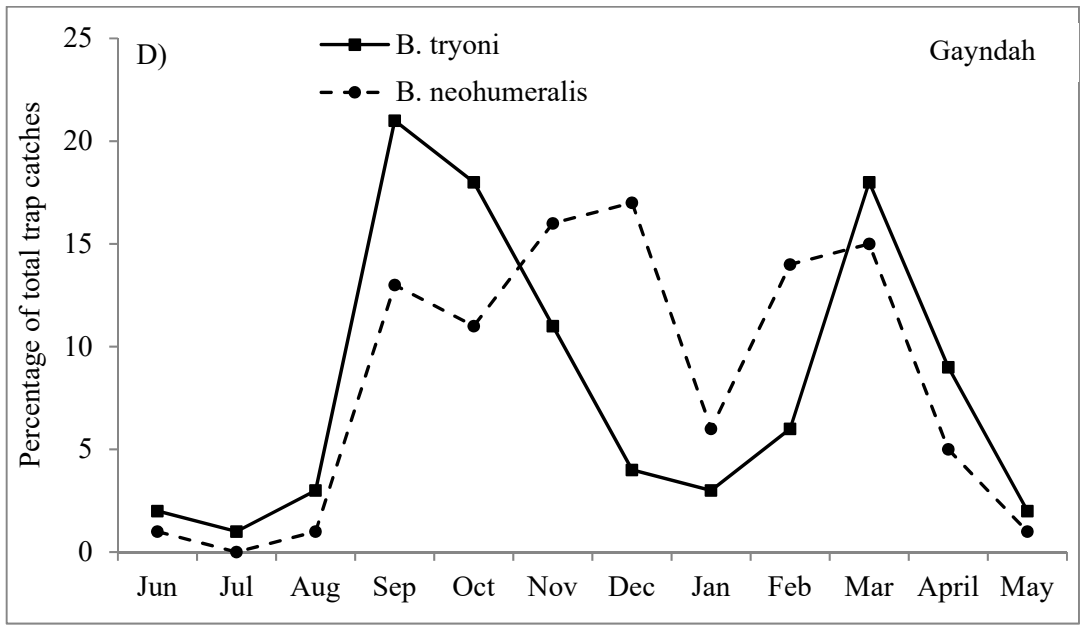


Figure 3. South Johnston

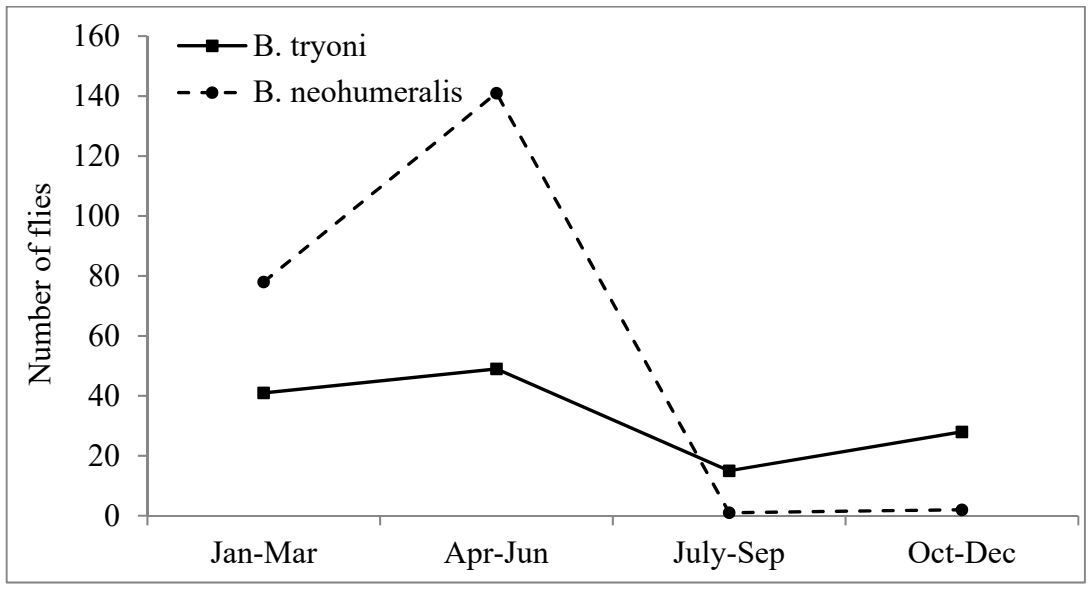
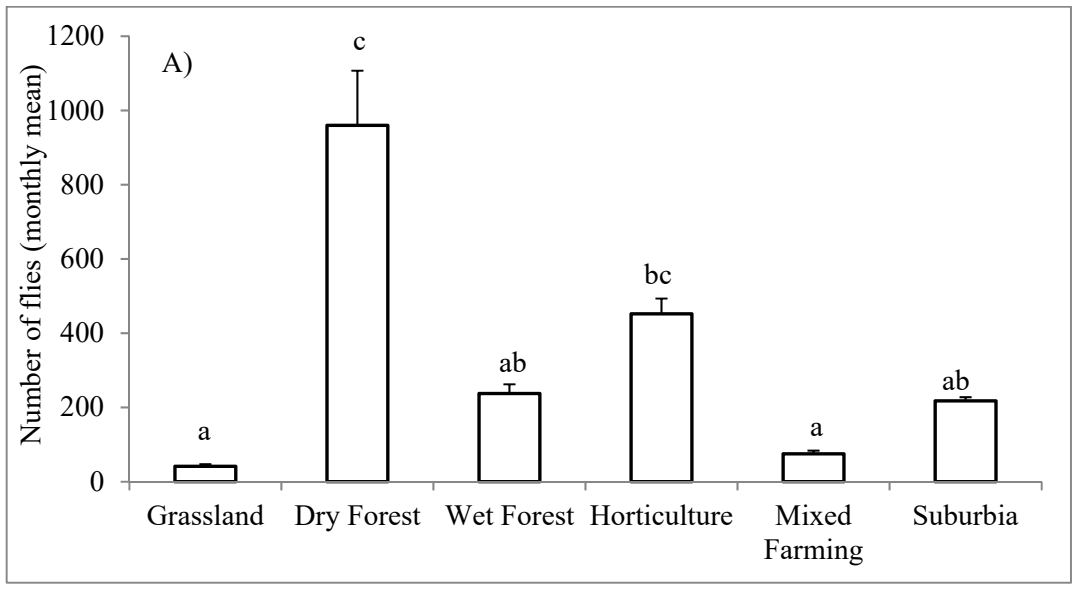


Figure 3. Abundance of (A) *Bactrocera tryoni* and (B) *Bactrocera neohumeralis* in different landscapes



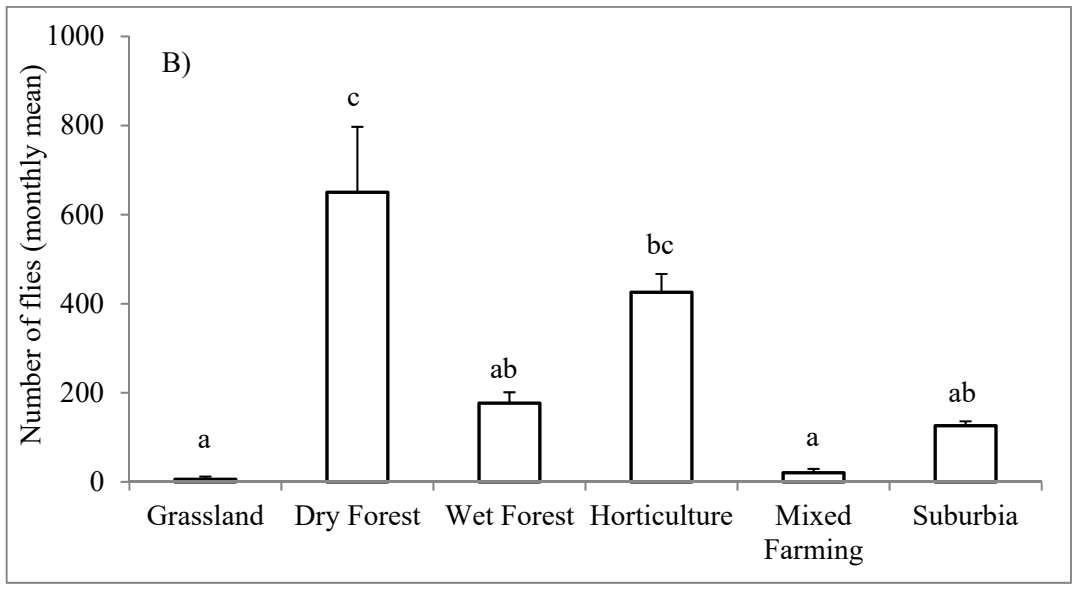


Figure 4. Comparison of abundance of *Bactrocera tryoni* and *Bactrocera neohumeralis* in different landscapes. Trap catches are mean number of flies per month collected over 7 months

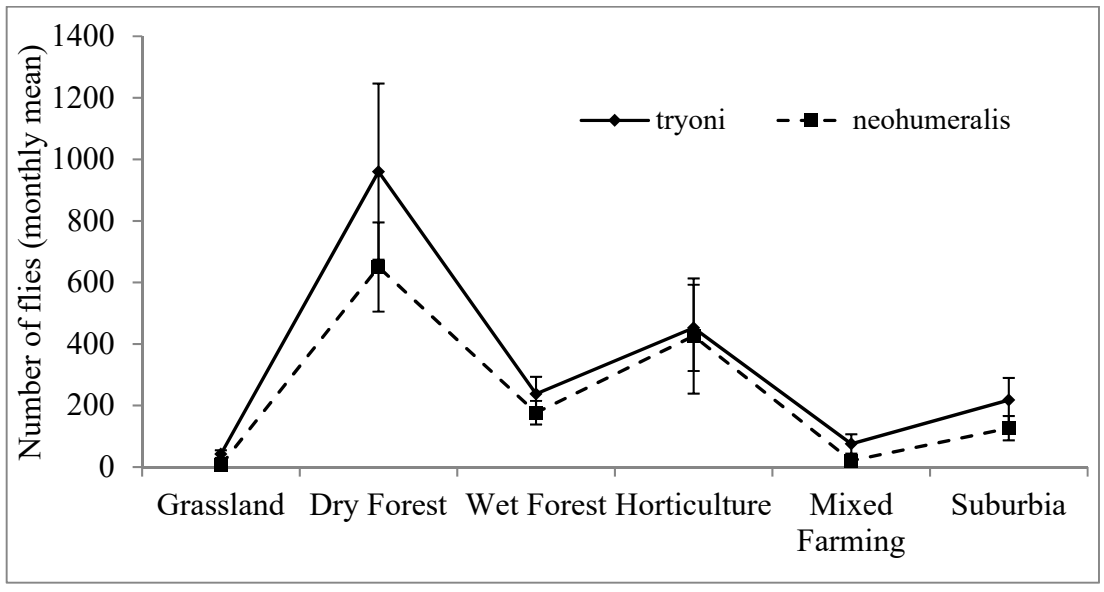
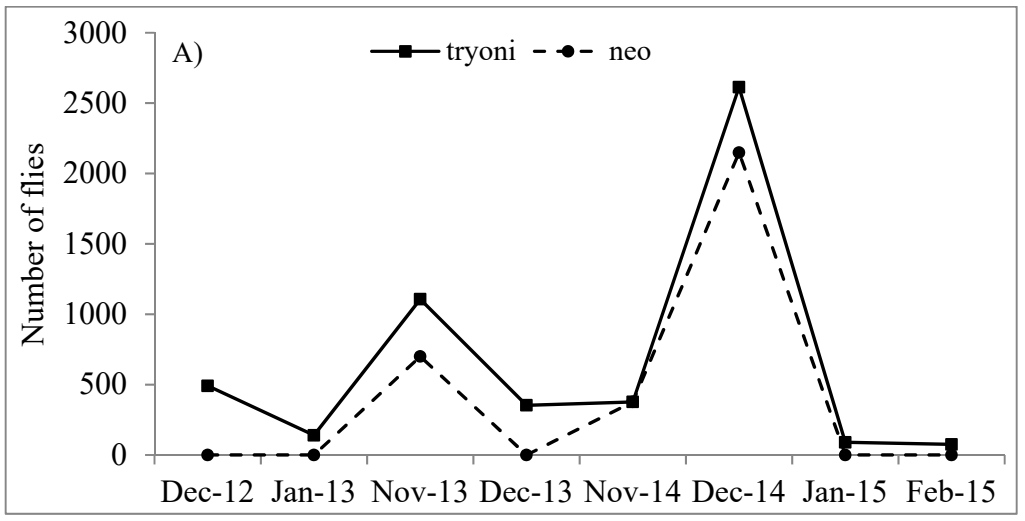
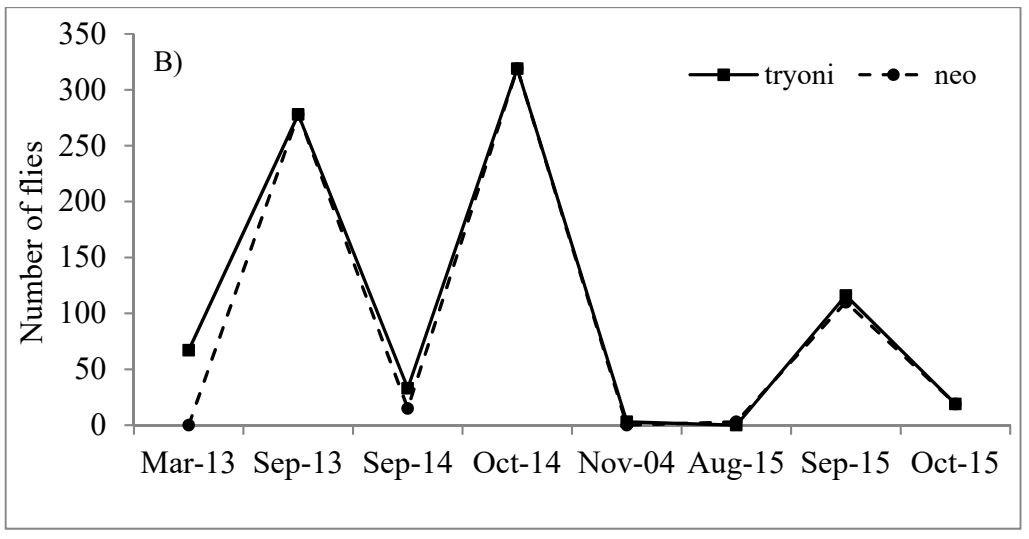


Figure 5a. Comparison of host use pattern between *Bactrocera tryoni* and *Bactrocera neohumeralis* A) White sapote B) Mulberry C) Plum





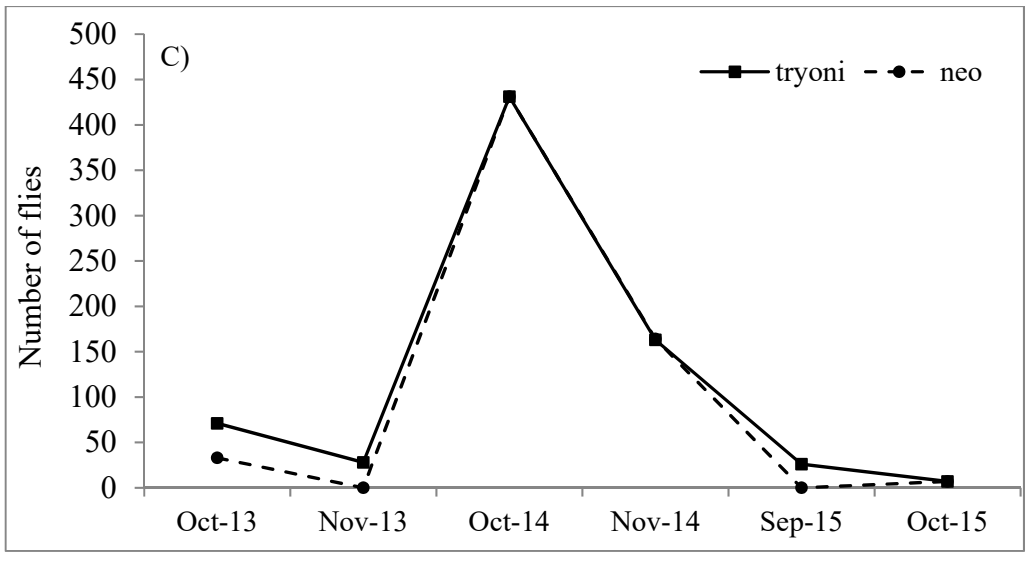
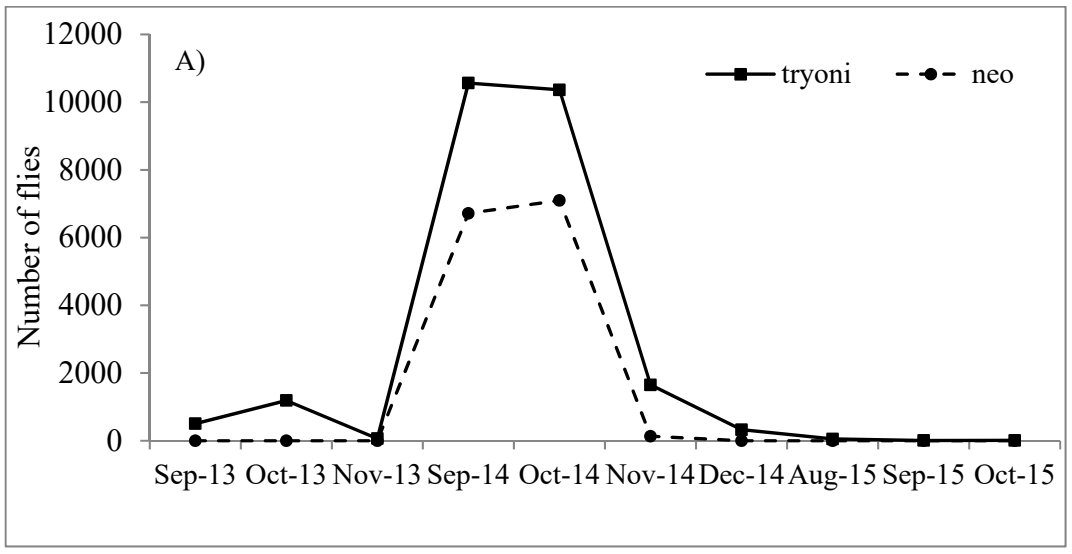
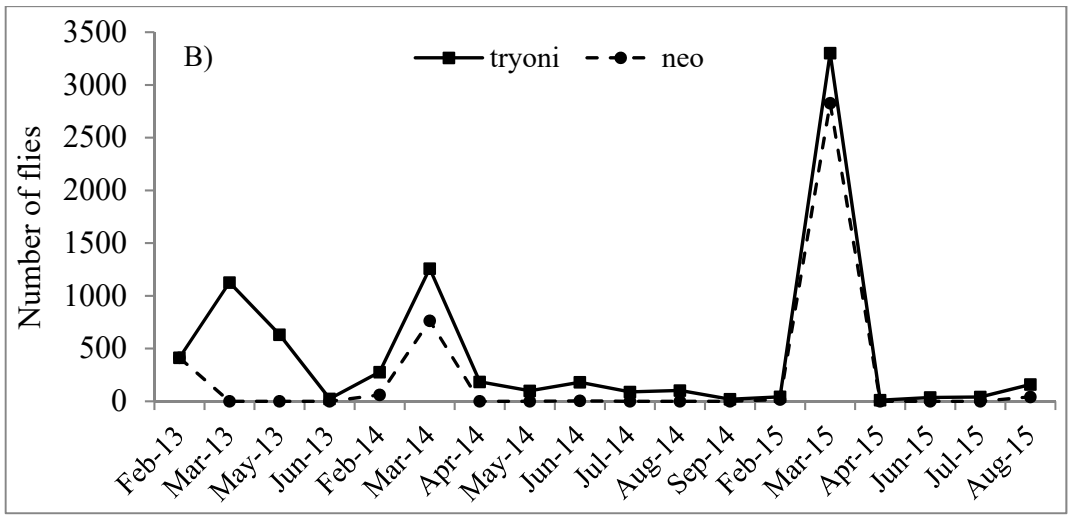


Figure 5b. Comparison of host use pattern between *Bactrocera tryoni* and *Bactrocera neohumeralis* A) Peach B) Carambola C) Nectarine





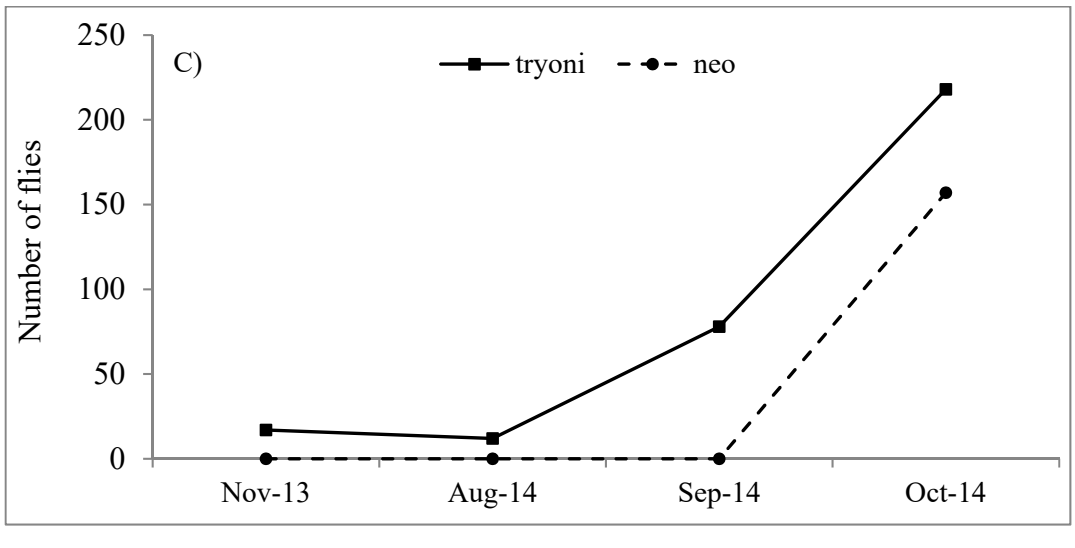


Figure 6. Abundance of *Bactrocera tryoni* and *Bactrocera neohumeralis* in different host fruits