Article Type: Special Issue Contribution

Article Title: Strong genetic differentiation but limited niche partitioning in a sympatric species pair separated by an allochronic reproductive barrier

Mitchell Irvine[1]*, Zachary Stewart[1], Nagalingam Kumaran[2], Chapa G. Manawaduge[1], James Ryan[1], Solomon Balagawi[3], Brendan Missenden[4], Melissa Starkie[4], Anthony R. Clarke[1], David Hurwood[1], and Peter Prentis[1]

Author Information

Correspondence author: Mitchell Irvine (ORCID ID – 0009-0004-2338-8463)

mitchell.irvine@hdr.qut.edu.au

[1]School of Biology and Environmental Science, Queensland University of Technology, GPO Box 2434, Brisbane, Qld, Australia

[2]Commonwealth Scientific and Industrial Research Organization (CSIRO), Ecosciences Precinct, Dutton Park, Queensland, Australia.

[3]Horticulture and Forestry Science, Department of Primary Industries, P.O. Box 652, Cairns, Queensland, Australia.

[4]Department of Primary Industries, Ecosciences Precinct, Dutton Park, Queensland, Australia

Running title: Genetic differentiation but limited niche partitioning

Abstract Word Count: 248

Word Count: 7,944 (excl. acknowledgements, author contributions, data accessibility and references)

Abstract

The combination of differential adaptation to ecological niches and the development of reproductive barriers are considered helpful for maintaining co-existing species. In the absence of one of these elements, species boundaries are expected to breakdown. The tephritid fruit fly species pair, *Bactrocera tryoni* and *B. neohumeralis*, have significant overlap in geographic range and host use, with time of day of male mating the only known difference in their mating systems. Using a combination of ecological (seasonal abundance, host use and habitat use) and genomic data, we tested the differing roles of competition and assortative mating on the maintenance of the species boundaries in this species pair. Genome-wide SNP analyses found strong genetic differentiation between the species with no evidence for hybridization in the field. Most outlier SNPs were restricted to narrow regions towards the centromeres and telomeres of chromosomes and high nucleotide diversity rates were observed throughout the chromosomes of both species. Enrichment of annotation terms indicated an overabundance of genes with the 'abnormal neuroanatomy' term. Terms of interest associated with sleep and circadian rhythm, potentially important to the allochronic reproductive barrier, were non-enriched. Ecological data found limited evidence for ecological differentiation between the two species based on significant positive correlations between species numbers trapped at different times of the year, trapped in different habitats within a region, or when reared from fruit.

Our study highlights the significance of assortative mating over ecological differentiation for sympatric maintenance of the *B. tryoni*/*B. neohumeralis* sibling pair.

Keywords

Speciation, *Bactrocera*, genomic divergence, niche partitioning, assortative mating, co-existence

Introduction

Understanding the ecological and genetic basis of speciation and species maintenance is a fundamental element of evolutionary biology. In geographically isolated populations, an absence of gene flow between incipient species is expected to allow for the accumulation of genetic divergence via both selection and drift (Feder et al., 2013). When two populations experience gene flow (either through primary or secondary contact), reproductive isolation is considered necessary to limit its homogenizing effect (Coyne & Orr, 2004).

Central to our understanding of speciation in the presence of gene flow are the opposing roles of natural selection and recombination on the formation or dissolution of species boundaries (Felsenstein, 1981). Natural selection promotes the development of reproductive isolation through the formation of locally adapted gene complexes, while free recombination acts to dissolve them through breaking down the associations of such complexes (Felsenstein, 1981; Hey, 2001; Schluter & Rieseberg, 2022); a reduction in recombination rates is considered helpful for species boundaries to persist and strengthen over time (Butlin, 2005). Traits associated with reproductive isolation, such as those involved in assortative mating, are considered effective mechanisms for constraining gene flow and are essential for sympatric species divergence when coupled with the development of traits associated with local adaptation (Ortiz-Barrientos et al., 2016; Schluter & Rieseberg, 2022).

On an ecological scale, the persistence of two species in sympatry depends on two criteria; their ability to overcome the effects of hybridization through the development of reproductive isolation, and the development of niche partitioning to avoid inter-species competition

asymmetries (Coyne & Orr, 2004; Felsenstein, 1981; Weber & Strauss, 2016). Without meeting these criteria, theory suggests that species will likely dissolve through hybridization or cease to exist in sympatry through competitive displacement (Germain et al., 2021; Weber & Strauss, 2016). Although the coexistence of ecologically equivalent species may occur long term where sexual selection exists, provided the local carrying capacity is sufficient enough to maintain both populations (M'Gonigle et al., 2012).

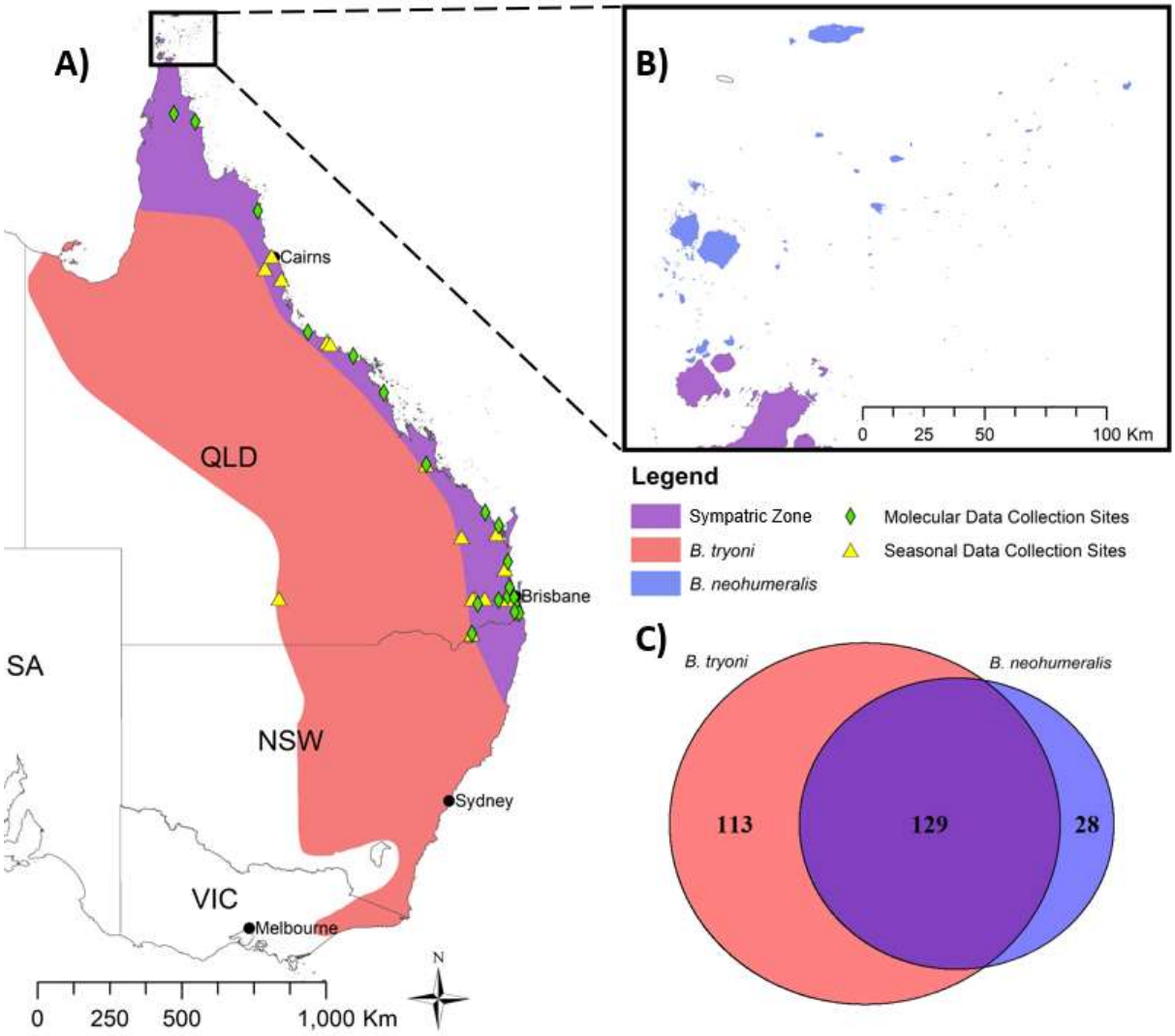
The Australian frugivorous tephritids (Diptera: Tephritidae), *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy), are a fruit fly sibling pair, studied for their speciation histories and genetic stability for over 60 years (Birch, 1961; Gilchrist et al., 2005; Gilchrist & Ling, 2006; Gilchrist et al., 2014; Lewontin & Birch, 1966; Meats et al., 2003; Pike, 2004; Raphael et al., 2019; Smith, 1979; Yeap et al., 2020). Genetically, *B. tryoni* and *B. neohumeralis* are very similar (Gilchrist et al., 2014; Morrow et al., 2000; Wang et al., 2003), with shared polymorphisms in coding and non-coding regions indicating continued genetic exchange or recent separation (Gilchrist et al., 2005; Gilchrist & Ling, 2006; Wang et al., 2003). The species show only minor variation in their morphology; the humeral calli, located on the 'shoulders' of the fly, are bright yellow in *B. tryoni* and brown in *B. neohumeralis* (Drew, 1989). While intermediate colour variants are well documented, genetic analysis using microsatellites suggest that intermediates represent intraspecific phenotypic variation, rather than hybrid individuals (Gilchrist & Ling, 2006; Pike, 2004; Wolda, 1967). While *B. tryoni* is typically more abundant than *B. neohumeralis*, the geographic ranges of the two species extensively overlap along the east coast of Australia, with the endemic range of *B. neohumeralis* entirely encompassed by the range of *B. tryoni* except for a few islands in the Torres Strait, north of Queensland (QLD) (Fig. 1A,B) (Osborne et al., 1997). Both *B. tryoni* and *B. neohumeralis* are highly polyphagous, using fruit from multiple plant families for breeding, with host utilization data

91 [presented as 'simple' lists of host fruit that either species have been reared from at least once

92 with little or no further information] showing significant overlap between the pair (Fig. 1C)

93 (Hancock et al., 2000; Leblanc et al., 2012; QDPC, 2021).

94



95

96 Figure 1. A) Distribution of *Bactrocera tryoni* and *B. neohumeralis* along the east coast of

97 Australia. B) Distribution of both species in the Cape York Peninsula and Torres Strait

98 Islands. C) Host usage overlap between species. Numbers in circles represent the known

99 number of host species utilized. Range and host data gathered from Hancock et al. (2000);

100 Leblanc et al. (2012); QDPC (2021). Note: Recent expansion during the past 50 years of *B. tryoni*

into central and western Victoria is not shown because it is still under regulatory control in many places and is not associated with the species' evolutionary history.

So closely related are *B. tryoni* and *B. neohumeralis* that only one known trait defines them as separate biological species: their time of mating (Birch & Vogt, 1970; Lewontin & Birch, 1966; Vogt, 1977). *Bactrocera neohumeralis* mates during the middle of the day at high light intensity over a mating window of three to seven hours, while *B. tryoni* initiates mating during dusk at low light intensity in a mating window of approximately one hour (Ekanayake et al., 2017). This difference in mating time is the only known mechanism maintaining reproductive isolation between the species (Birch & Vogt, 1970; Lewontin & Birch, 1966; Meats et al., 2003; Raphael et al., 2019; Smith, 1979), with other components of their mating system not known to be different (Bellas & Fletcher, 1979; Ekanayake et al., 2017; Mankin et al., 2008), including an absence of post-zygotic barriers (Yeap et al., 2020). Female *B. neohumeralis* will, like *B. tryoni*, also mate at dusk, at which time *B. tryoni* males do not discriminate between females of either species (Yeap et al., 2020). This caused Yeap et al. (2020) to conclude "*the mating time difference may be a weaker reproductive isolating barrier than once assumed*".

If mating time is a weak reproductive barrier between the two species, then other factors must be playing a role in maintaining separate species; for example, genomic incompatibility (Dobzhansky, 1936; Turelli et al., 2001). Additionally, for herbivorous insects such as fruit flies, species divergence is largely considered to be driven by host switching and host specialization that reduces the effects of competition and natural enemies (Gavrilets & Losos, 2009; Janz & Nylin, 2008; Jeffries & Lawton, 1984; Schluter, 2000; Yoder et al., 2010) and so ecological factors also need to be considered when investigating the maintenance of species boundaries. Conversely the species may not be evolutionarily stable, in which case evidence of gene flow should be expected from sympatric field collections of the two taxa.

125 In this study, we conducted a comprehensive investigation of how the lineages of *B. tryoni*
126 and *B. neohumeralis* species have diverged and been maintained in apparent sympatry
127 through dense geographic sampling of nuclear genomes, and structured field sampling of
128 temporal population abundance and habitat- and host-usage. We assessed the extent of
129 ongoing gene flow and genetic differentiation between the species and identified putative
130 candidate genes associated with reproductive isolation. Using the field data, we conducted
131 correlation analyses on seasonal abundance, host use and habitat utilization for evidence of
132 competitive displacement between the two species. Shifts in any or all of these three
133 ecological attributes have been previously associated with competition between fruit fly
134 species (Clarke & Measham, 2022).

135 **Methods**

136 *Molecular data sample collection*

137 Individuals of *B. tryoni* and *B. neohumeralis* were collected in December 2020 and December
138 2021, using cue lure baited Steiner (Steiner, 1957), Lynfield (Cowley et al., 1990) or Paton traps
139 (Plant Health Australia, 2023). These traps were deployed by Queensland Department of
140 Primary Industries (QDPI) as part of ongoing Biosecurity Queensland (BQ) surveillance. A
141 subsample of the trapping sites was selected from the existing network to represent the
142 geographic range of both species (Fig. 1A). Up to 10 individuals of each species were
143 collected per site with a total of 81 *B. tryoni* and 105 *B. neohumeralis* individuals collected
144 from 19 sites. Detailed trapping records are included in Table S1. Identity of the samples were
145 confirmed following the descriptions in fruit fly identification handbooks (Drew, 1989; Plant
146 Health Australia, 2018), and samples were stored at -20°C in 100% ethanol.

147 *Extraction and sequencing*

148 DNA was extracted using the QIAGEN DNeasy® Blood & Tissue Kit following the
149 manufacturer's protocol. Samples were screened for quality and quantity, using both gel
150 electrophoresis and Qubit assay, and then sent to Diversity Arrays Technology Pty Ltd,
151 Canberra (DArT P/L), for DArTseq high-density genotyping (Kilian et al., 2012). A PstI/SphI,
152 restriction enzyme combination was used by DArT P/L, and the fragments (up to 90bp) were
153 sequenced on an Illumina Hiseq2500 as single end reads.

154 *Data sourcing*

155 All relevant analyses made use of the *B. tryoni* genome assembly (Genbank accession =
156 GCF_016617805.1).

157 *Read alignment and SNP calling*

158 Reads were demultiplexed using process_radtags in Stacks v2.60 (Rochette et al., 2019). Read
159 mapping used BWA-MEM (Li, 2013) to produce a SAM alignment file with appropriate read
160 group specification for each sample. Subsequently, SAMtools (Li et al., 2009) produced sorted
161 and indexed BAM alignment files. Freebayes v1.3.6 (Garrison & Marth, 2012) was used in an
162 iterative process to predict single nucleotide polymorphisms (SNPs) in all samples. Firstly,
163 each sample had SNP calling performed individually using freebayes; output Variant Call
164 Format (VCF) files were left-aligned using BCFtools (Danecek et al., 2021) and block
165 substitutions were decomposed using vt (Tan et al., 2015). Resulting VCF files were merged
166 using BCFtools, and this VCF file was then used as input for a second round of SNP calling
167 using freebayes (*-@ file.vcf --only-use-input-alleles*) to ensure all samples were genotyped at
168 the same locations. The variant calls were filtered using a method based upon James et al.
169 (2021) wherein SNPs were removed if they had missing data in > 50% of the population, had
170 a quality score < 30, a depth < 3 read alignments, or a minor allele count < 1; we also filtered
171 out SNPs with a minor allele frequency < 5%.

Population structure and genetic differentiation

To investigate whether admixture was occurring between the two species, we used fastSTRUCTURE v1.0 (Raj et al., 2014). A simple prior of K=1-10 was used, after which the chooseK.py utility program provided with fastSTRUCTURE was used to assess the optimal number of genetic clusters to explain the structure among species and populations. Plots were generated using pophelper v2.3.1 (Francis, 2017) in R.

Additionally, we ran a Principal Component Analysis (PCA) to provide further visualization of the variations within and between the two species, using the SNPRelate v1.24.0 (Zheng et al., 2012) package in R. The linkage disequilibrium-based SNP pruning function provided with this software was run with default parameters to mitigate any bias associated with linked SNP clusters upon the downstream PCA. We similarly performed this SNP pruning prior to calculation of IBS proportions for samples using SNPRelate.

Relatedness prediction

Relatedness analyses were performed in R using the dartRverse package (Gruber et al., 2018; Mijangos et al., 2022). Variants were filtered to remove those with a callrate lower than 95%, those deviating from Hardy-Weinberg Equilibrium to a significance threshold of 0.05, and those with linkage disequilibrium values above 0.2. The remaining variants were used as input for the EMIBD9 (Wang, 2022) and rrBLUP (Endelman, 2011) packages to estimate the relatedness of samples in a pairwise manner.

Population genetics statistics

Mean F_{ST} (θ) values (Weir & Cockerham, 1984) were calculated both within and between the two species using VCFtools (Danecek et al., 2011). Within species comparisons were conducted for each species separately and samples from each collection site were considered

separate populations. For between species comparison, the samples were grouped by species and entered as two separate populations.

Additional population statistics were calculated to estimate levels of relative genetic differentiation between species (F_{ST}), absolute genetic divergence between species (dXY), and nucleotide diversity within species (π). These statistics were calculated within non-overlapping windows of 50 Kbp using pixy (Korunes & Samuk, 2021). As pixy explicitly requires invariant sites to be called, we recalled variant and invariant sites using a pipeline involving bcftools mpileup (-p -A -d 500) followed by bcftools call (-m), with resulting predictions used as input. These statistics were mapped along each chromosome to identify patterns of differentiation within the chromosomal landscape.

Recombination rate prediction

The recombination rate along each chromosome was approximated using a recurrent neural network as implemented by ReLERNN (Adrion et al., 2020). The variants for each species were filtered to retain only biallelic SNPs, after which each species' data was processed using ReLERNN_SIMULATE to generate training and validation data (with default settings), ReLERNN_TRAIN to train the neural network (with --nEpochs 100 and --nValSteps 20), ReLERNN_PREDICT to predict per-base recombination rate, and ReLERNN_BSCORRECT to perform parametric bootstrapping and estimate 95% confidence intervals.

Population structure along chromosomes

Population structure was assessed within local regions of each chromosome as implemented by the LOSTRUCT method (Li & Ralph, 2018). The run_lostruct.R wrapper script (available within the program's GitHub repository at https://github.com/petrelharp/local_pca) was configured to assess local structure using PCA within non-overlapping windows of 50 SNPs. For each window, the first principal component was extracted for each sample and the spread of these

219 values were visualized by their median and the 25% and 75% quantiles to identify regions
220 where population structure was more apparent.

221 *Outlier SNPs and gene candidates*

222 Variant predictions were converted from VCF to GESTE format using PGDSpider v2.1.1.5
223 (Lischer & Excoffier, 2012), then outlier SNPs were predicted using Bayescan v2.1 (Foll &
224 Gaggiotti, 2008) with default parameters. Genes containing outlier SNPs within their gene
225 model (exons and/or introns), as well as genes that were the nearest to an intergenic outlier
226 SNP were identified using a custom script.

227 Candidate genes were queried against the *Drosophila melanogaster* gene models (Larkin et al.,
228 2021) (FlyBase release FB2023_01) using MMseqs2, and the FlyBase automated gene
229 summaries of the best hit were parsed for functional and phenotypic annotations which we
230 attributed to sequences if a minimum E-value of $1e^{-10}$ was achieved.

231 *Seasonal abundance*

232 A historical dataset (May, 1961) was used for the comparison of temporal abundance of *B.*
233 *neohumeralis* with *B. tryoni*. This dataset has been previously used to report *B. tryoni* [but not
234 *B. neohumeralis*] seasonal phenology (Muthuthantri et al., 2010), and so the phenology patterns
235 *per se,* are not explored here. Rather, we focus on evidence for temporal niche segregation in
236 *B. neohumeralis* and *B. tryoni* through paired t-tests and Pearson correlation analysis using
237 the R package ggpubr (Kassambara, 2023). In Africa, a shift in seasonal abundance to
238 minimize an overlap in host usage (i.e. temporal niche displacement) has been reported as the
239 outcome of competition between the native mango fruit fly (*Ceratitis cosyra* (Walker)) and
240 invasive Oriental fruit fly (*Bactrocera dorsalis* (Hendel))(Vayssières et al., 2015). Therefore,
241 we predict that if there is competition-driven temporal niche displacement between the two

species, then there should be no, or even negative correlation between their seasonal abundances. We interpret significant positive correlation between their populations as a lack of evidence of temporal niche displacement.

Detailed methodology on the fly trapping dataset is described elsewhere (May, 1961; Muthuthantri et al., 2010). In summary, 15 trapping sites across Queensland, Australia, were used, ranging from temperate Stanthorpe in the south to tropical Cairns, ~1500 km to the north (see Fig.1A and Table S5). Fly trapping was carried out for up to six years per location using a fruit-based liquid lure trap (May & Caldwell, 1944). Ten traps were maintained at each location, with traps cleared weekly. The data in May (1961) vary between sites in how they are presented, and modifications were needed in order to directly compare *B. tryoni* and *B. neohumeralis* populations. Six sites were excluded as seasonal data were not available (see Table S5 for details). The t-tests and correlation analyses between *B. tryoni* and *B. neohumeralis* populations were conducted on a site-by-site basis (i.e., data were never collated across sites), so the different manipulations performed on the site-specific data sets did not affect results.

Given the variability in how data from different locations are presented in the historical data, a dataset systematically collected as part of a recent landscape genetics study was also used to assess seasonal abundance on a fine scale. Detailed sampling techniques are presented in Ryan et al. (In Review) with abundance data being used for the seasonal abundance correlation analysis and the landscape ecology analysis below. In brief, samples were collected from 25 locations in the Wide Bay-Burnett Region of Queensland (see Fig. 2) in April 2021, August 2021, October 2021, December 2021, February 2022 and April 2022. The t-tests and correlation analyses between *B. tryoni* and *B. neohumeralis* populations was conducted for each bi-monthly collection.

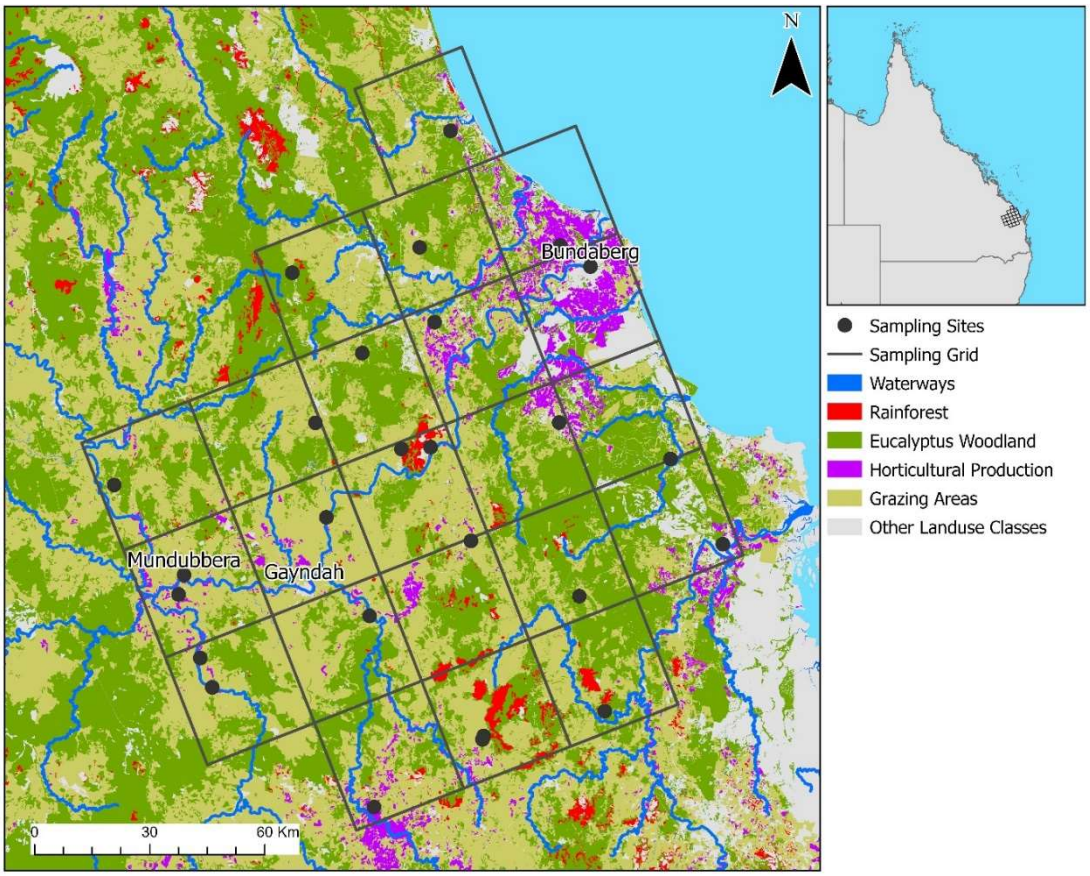


Figure 2: Collection sites throughout the Wide Bay-Burnett region of South East Queensland for the fine scale seasonal abundance analysis. Taken from Ryan et al. (In Review).

Habitat use

This study compared abundance of *B. tryoni* and *B. neohumeralis* across different human-defined habitat types to determine habitat use patterns. Six habitat types, grassland, suburbia, horticultural farming, mixed sugarcane and other crops (mixed farming hereafter), dry forest, and wet forest were chosen for the study in the Bundaberg region of Queensland (~350 km north of Brisbane). Specific descriptions of each habitat type are outlined in Table S7. For each habitat type there were nine replicate fruit fly traps. Trapping commenced in September 2010 (spring) and was completed in March 2011 using cuelure baited modified Steiner traps (Drew et al., 1982). The trapping period covers the period of peak yearly activity for *B. tryoni*

(Clarke et al., 2022). 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 were counted and identified to species level using a fruit fly identification guide (Drew, 1989). Paired t-tests and Pearson correlation analyses on the total trapping dataset for a habitat type were performed to compare the relationship in abundance of the two different species at that habitat type. Examples of spatial displacement of competing fruit flies are presented in Duyck et al. (2004).

Landscape ecology

Linear mixed modelling (LMM) was used to investigate patterns in *B. neohumeralis* abundance across the study area (See Fig. 2) to complement a study conducted by Ryan et al. (In Review) which investigated the environmental predictors of *B. tryoni* abundance. Additionally, the Empirical Bayesian Kriging (Geostastical Analyst Tools) tool was implemented in ArcGIS Pro v3.4 to interpolate *B. neohumeralis* abundance across the study area to visualise spatially explicit patterns of abundance.

The *B. neohumeralis* abundance data collected throughout the Wide Bay-Burnett region described in the fine scale seasonal abundance analysis above was used. The LMM analysis incorporated the same landscape predictor variables as was used in the LMM analysis in Ryan et al. (In Review), briefly, % coverage of five landscape categories (grazing area, residential area, *Eucalyptus* woodland, rainforest and agricultural production) within a 1km radius of each sampling site were investigated. Additionally, other variables investigated included latitude, longitude, and distance from each sampling site to the nearest perennial water source.

Based on the LMM analysis methods outlined in Ryan et al. (In Review), the R package lme4 version 1.1-30 (Bates et al., 2015) was used to create our LMMs and 'dredge()' function implemented in the R package MuMIn version 1.47.1 (Barton, 2022) to assess all combinations of predictor variables, using sampling time as a random variable. Following dredging, LMMs with an ΔAICc < 2 were used to perform model averaging and determine for each landscape predictor; linear coefficient, standard error, statistical significance and upper and lower 95% confidence intervals.

Host fruit use

Fruits were collected from the QDAF Maroochy Research Facility, Nambour, Queensland (~100 km north of Brisbane) from separate orchard blocks that consisted of either mixed tropical 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. Fruits collected were: white sapote (*Casimiroa edulis*), mulberry (*Morus nigra*), peach (*Prunus persica*), plum (*Prunus domestica*), carambola (*Averrhoa carambola*), nectarine (*Prunus persica*), guava (*Psidium guajava*), sapodilla (*Manilkara zapota*), *Syzigium* spp., grumichama (*Eugenia brasiliensis*), hog plum (*Spondias mombin*), jabotica (*Plinia cauliflora*), white mulberry (*Morus alba*), avocado (*Persea americana*), black sapote (*Diospyros nigra*), longan (*Dimocarpus longan*), mango (*Mangifera indica*) and cashew (*Anacardium occidentale*).

Fruit were placed in paper bags and transported to the QDAF Brisbane 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

324 Environment Room (26°C and 70%RH) to allow insects to develop through to the pupal
325 stage. The vermiculite was sieved weekly until all insects had exited fruit and pupated, and
326 fruit were inspected before being discarded. Fruit fly pupae that were reared out of fruit
327 samples were placed into small plastic boxes with gauzed lids containing vermiculite. Once
328 adult fruit flies had emerged from pupation, they were identified to species level using
329 morphological taxonomic characters. The abundance of the two fly species from a host was
330 compared using a paired t-test and Pearson correlation analysis. Changes in fruit fly host
331 usage patterns due to competition are well documented in the fruit fly literature (Charlery de la
332 Masselière et al., 2017; Hassani et al., 2022).

333 **Results**

334 Molecular Data

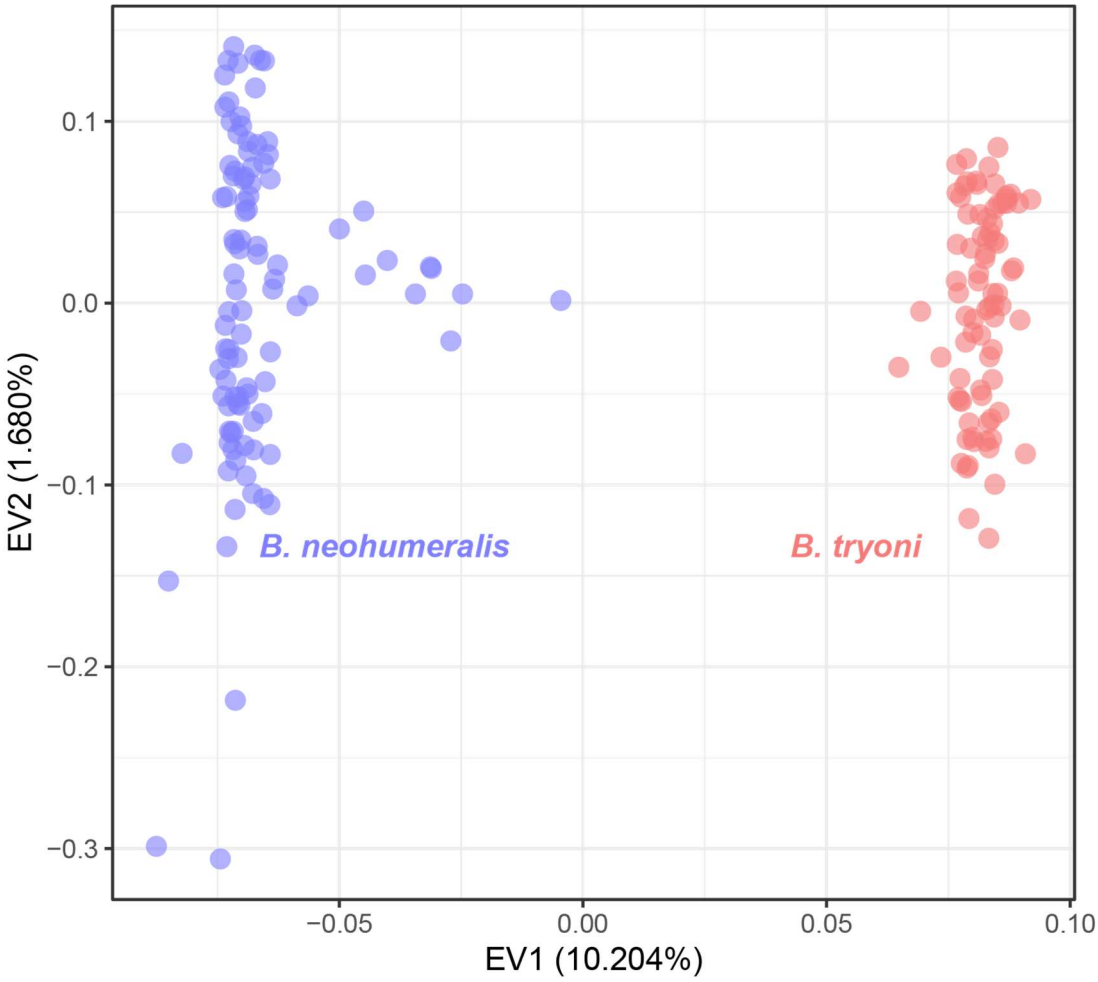
335 *Read processing, alignment and SNP calling*

336 After demultiplexing samples, between 2,596,430 and 5,228,647 high-quality sequence reads
337 were obtained for each sample. Alignment to the *B. tryoni* genome assembly using BWA-
338 MEM indicated that most samples had high mapping rates (median alignment percentage =
339 74.98%), with a small number of samples having low sequence alignment (minimum
340 alignment percentage = 0.38%); see Table S2 and Table S3 for full details of read counts and
341 alignment percentages. 28,334 SNPs were identified using Freebayes from the two species
342 that met all predetermined threshold criteria for inclusion.

343 *Populations cluster by species and not geography*

344 We found that populations clustered based on species and not geography in all analyses. PCA
345 (Fig. 3) and fastSTRUCTURE (Fig. 4) both resolved species-specific groups, demonstrating
346 clear genetic delineation of the two species and an absence of intermediate genotypes between
347 the species. Several *B. neohumeralis* samples (genotype codes 526, 528, 529, 530, and 531

348 from the NS location) appear to be outliers based on PCA (See Fig. S1), but one sample from
349 the same location (527) grouped with the majority of *B. neohumeralis* samples in ordination
350 space. Additionally, the chooseK.py function of fastSTRUCTURE predicted three population
351 clusters, with the outlier samples from the PCA representing the third cluster (See Fig. S2).
352 These outlier samples had poor genotype calling rates (min = 27.9%, max = 32.3%) compared
353 to sample 527 (91.3%) and most other samples (median = 89.1%). As such, we believe these
354 outliers to result from sequencing issues and were excluded from further analysis.



355

356 Figure 3. Principal component analysis (PCA) of single nucleotide polymorphisms (SNPs)
357 *Bactrocera neohumeralis* and *B. tryoni*, with the values from the first two eigenvectors

plotted for each individual. Eigenvectors one (horizontal) and two (vertical) explain 10.204% and 1.68% of the variance in the data, respectively.

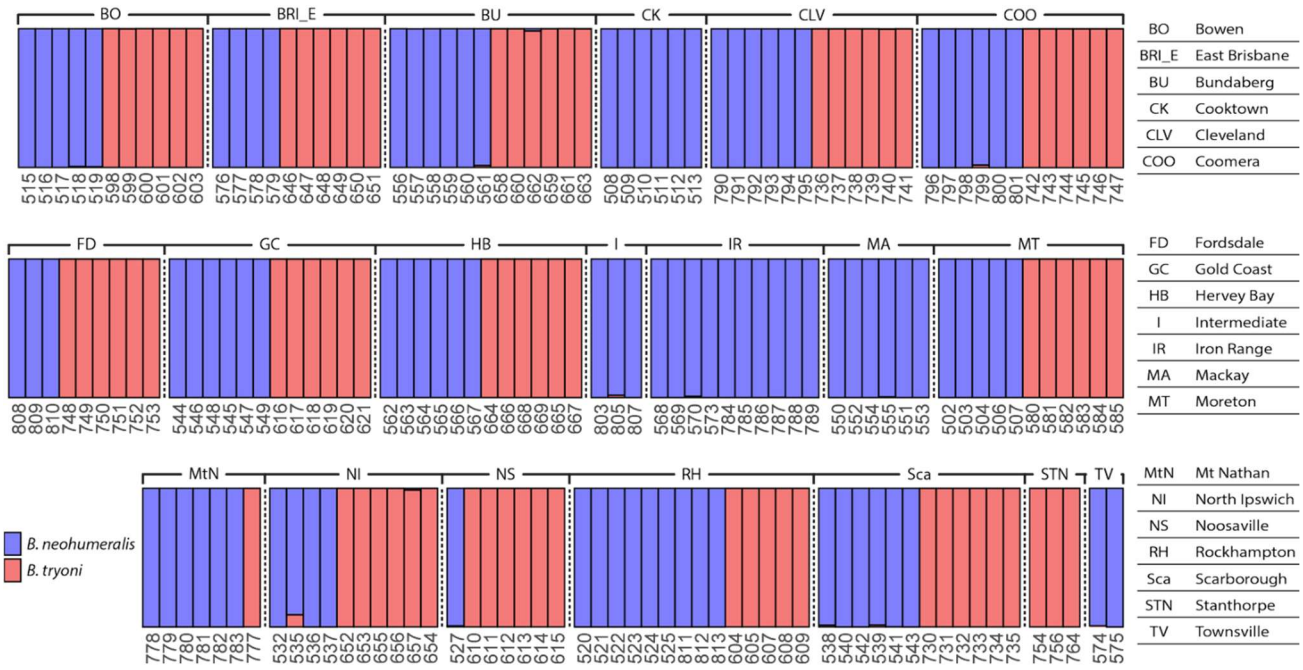


Figure 4. fastSTRUCTURE results at cluster size (K) equals two for single nucleotide polymorphisms (SNPs) for *Bactrocera neohumeralis* and *B. tryoni* grouped by collection site. Red and blue colouration of the bars indicates the inferred population admixture proportions as a percentage (numbers not shown). Three-digit numbers along the bottoms of each row of bars are genotype ID codes. The intermediate grouping (I) includes samples with an intermediate phenotype. All three intermediate samples (genotypes 803, 805 and 807) presented as *B. neohumeralis* genotypes.

Relatedness of individuals

After filtering, 3710 variants remained for use in the relatedness analyses. The EMIBD9 estimates of relatedness (i.e., Identity by Descent) for the sampled individuals are presented in Figure 5. Notably, each species is largely segregated into distinct groups showing increased relatedness values within species groups compared to between species groups where values

approached zero. Three samples (genotype codes 662, 787, and 788) appear as an outgroup relative to the other samples and show elevated levels of relatedness both within the same species and between species. While the levels of relatedness are elevated in these samples, they still present higher levels of relatedness within species groups compared to between species groups. These samples do not appear as anomalies in any other analysis suggesting a lack of data after filtering preventing the EMIBD9 analysis from discerning these samples resulting in elevated levels of relatedness.

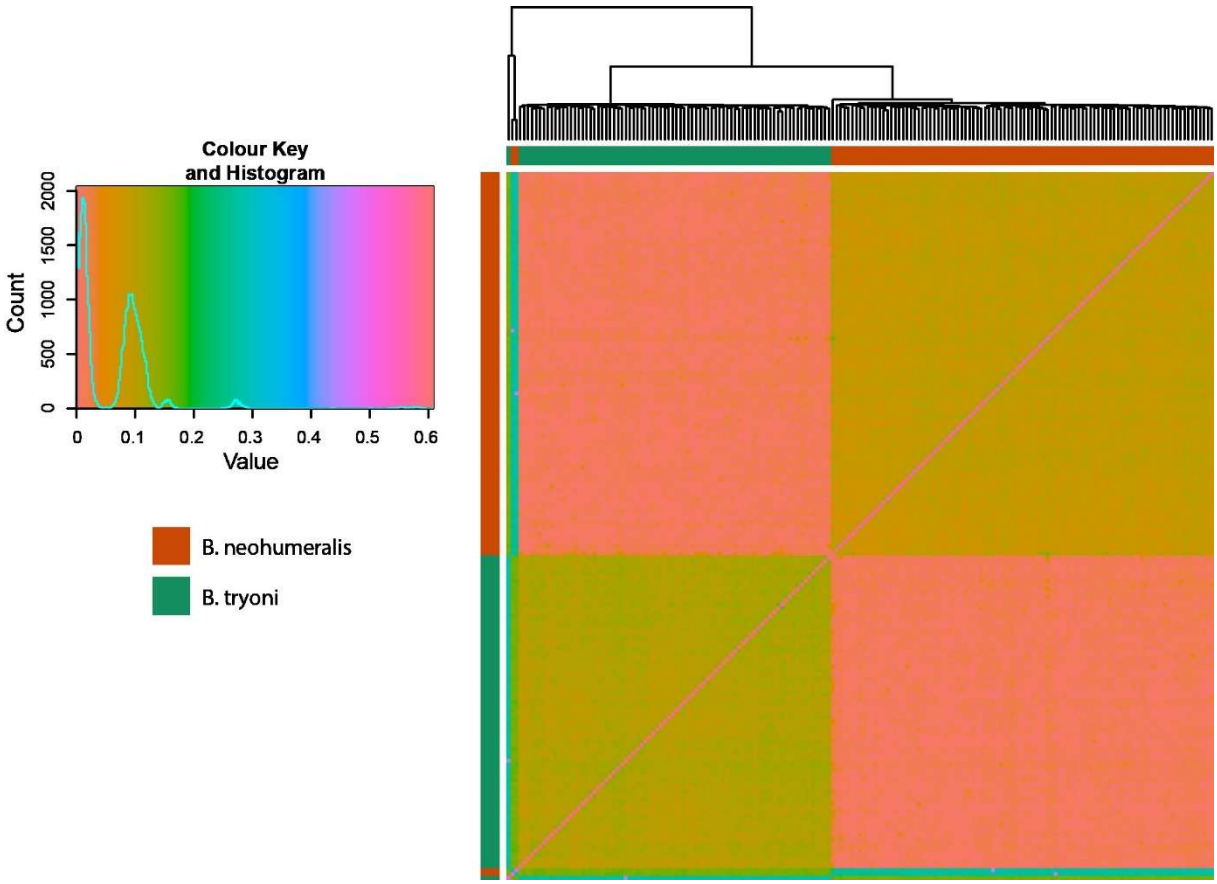
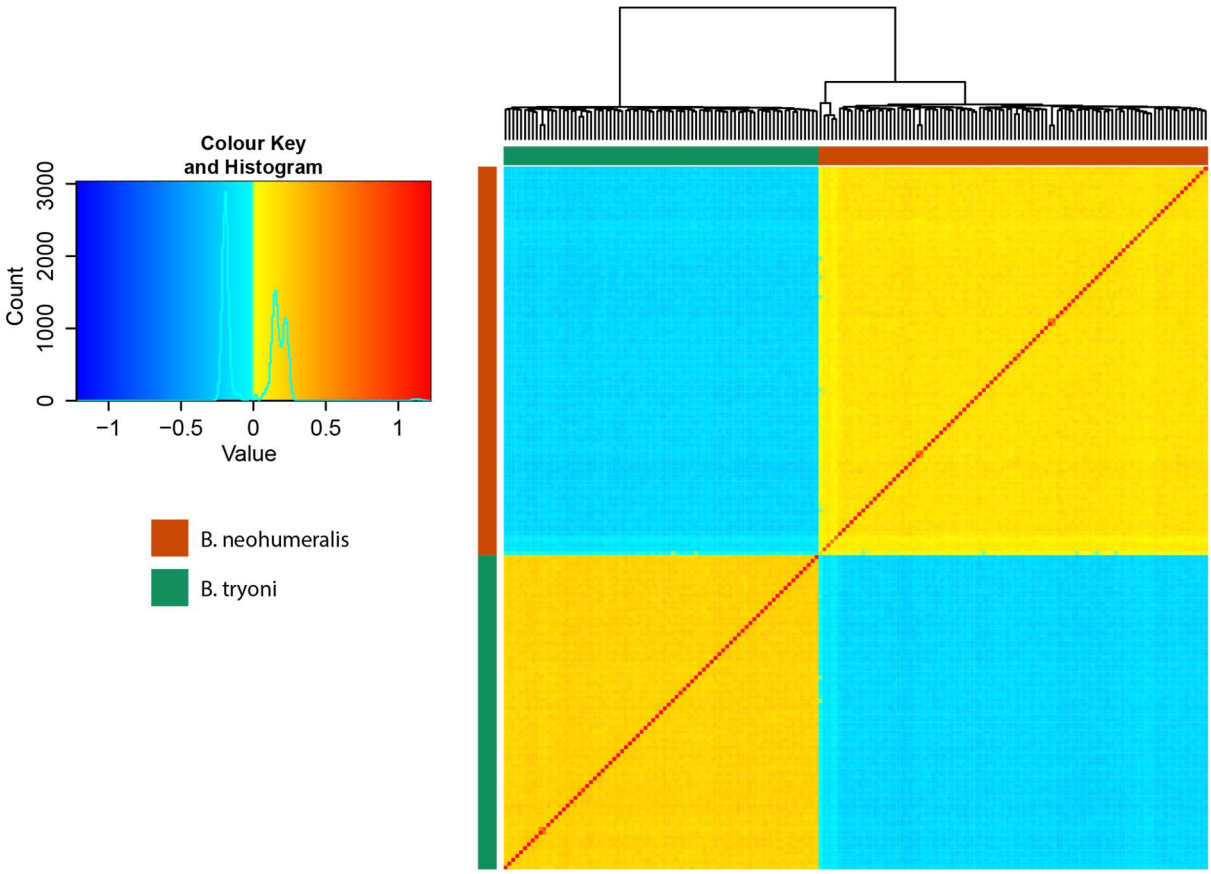


Figure 5. Heatmap generated using the results of EMIBD9 assessing the relatedness of sampled individuals via their Identity by Descent (IBD) coefficients. The colour key indicates the coefficient value with a histogram plotted on top of the colour key to showcase the amount of pairwise sample comparisons obtaining that value. The heatmap indicates the IBD

value for each pairwise sample comparison, with the sample's species indicated along the left and top of the plot. Lower values (e.g., soft red colour) indicate low to no relatedness as seen when comparing most *Bactrocera neohumeralis* samples against most *B. tryoni* samples. An increased value (e.g., amber colour) indicates some degree of relatedness, and is seen when comparing most species amongst themselves.

The rrBLUP estimates of relatedness are presented in Figure 6. Notably, genotype codes 662, 787, and 788 do not appear to be outliers in this analysis unlike what was seen in the results of EMIBD9; all species only show relatedness to other members of their species group with negative values obtained when compared to the opposite species suggesting non-random mating between the species groups.



396 Figure 6. Plot generated using the results of rrBLUP assessing the estimated relatedness of
397 sampled individuals via an additive relationship matrix. The colour key indicates the
398 probability of identity by descent with a histogram plotted on top of the colour key to
399 showcase the amount of pairwise sample comparisons obtaining that value. A heatmap
400 indicates the estimated relatedness value for each pairwise sample comparison, with the
401 sample's species indicated along the left and top of the plot. Negative values (e.g., blue
402 colour) indicate low to no relatedness as seen when comparing *Bactrocera neohumeralis*
403 samples against *B. tryoni* samples. A positive value (e.g., yellow colour) indicates some
404 degree of relatedness and is seen when comparing individuals within a species.

405 *Population genetics statistics, recombination rates and structure along chromosomes*

406 Patterns of differentiation among populations (F_{ST}) were also lower between populations
407 within a species (*B. neohumeralis* = 0.0025796 when ignoring outlier samples, *B. tryoni* =
408 0.0036269) and two orders of magnitude higher between species (0.10843). Using outlier
409 detection in BayeScan, 357 of the 28,334 SNPs (1.26%) displayed a significantly elevated
410 level of differentiation between the species. The distribution of outlier SNPs along autosomal
411 chromosomes (See Fig. 7 to Fig. 11) was largely restricted to ~ 16 narrow genomic regions.
412 Elevated windowed F_{ST} and d_{XY} values calculated using Pixy were also clustered in similar
413 regions to those observed in the outlier SNPs. These narrow genomic regions with outlier loci
414 were not uniformly distributed across the five autosomal chromosomes with 10 occurring on
415 chromosome ends and two near the centre of chromosomes. A relatively small number of
416 outlier SNPs were mapped to 'unplaced contigs' (See Fig. S3). Identity-By-State (IBS)
417 analysis also provided insight into the genetic homogeneity of the two species, with *B. tryoni*
418 having a mean value of 0.797 and *B. neohumeralis* having a mean value of 0.790 (species

419 with a value closer to 1 are more similar) suggesting that both species are approximately

420 equally homogenous.

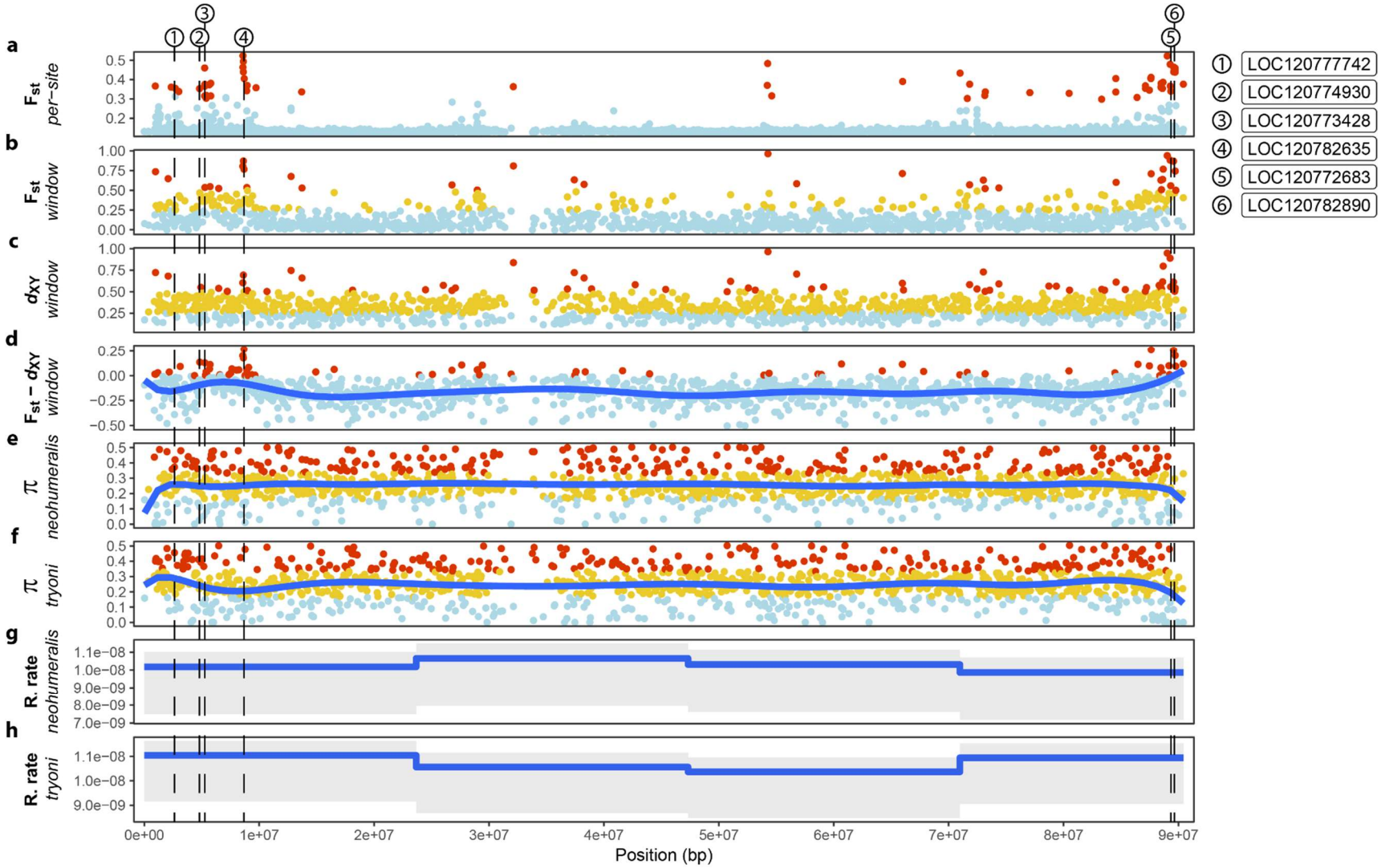
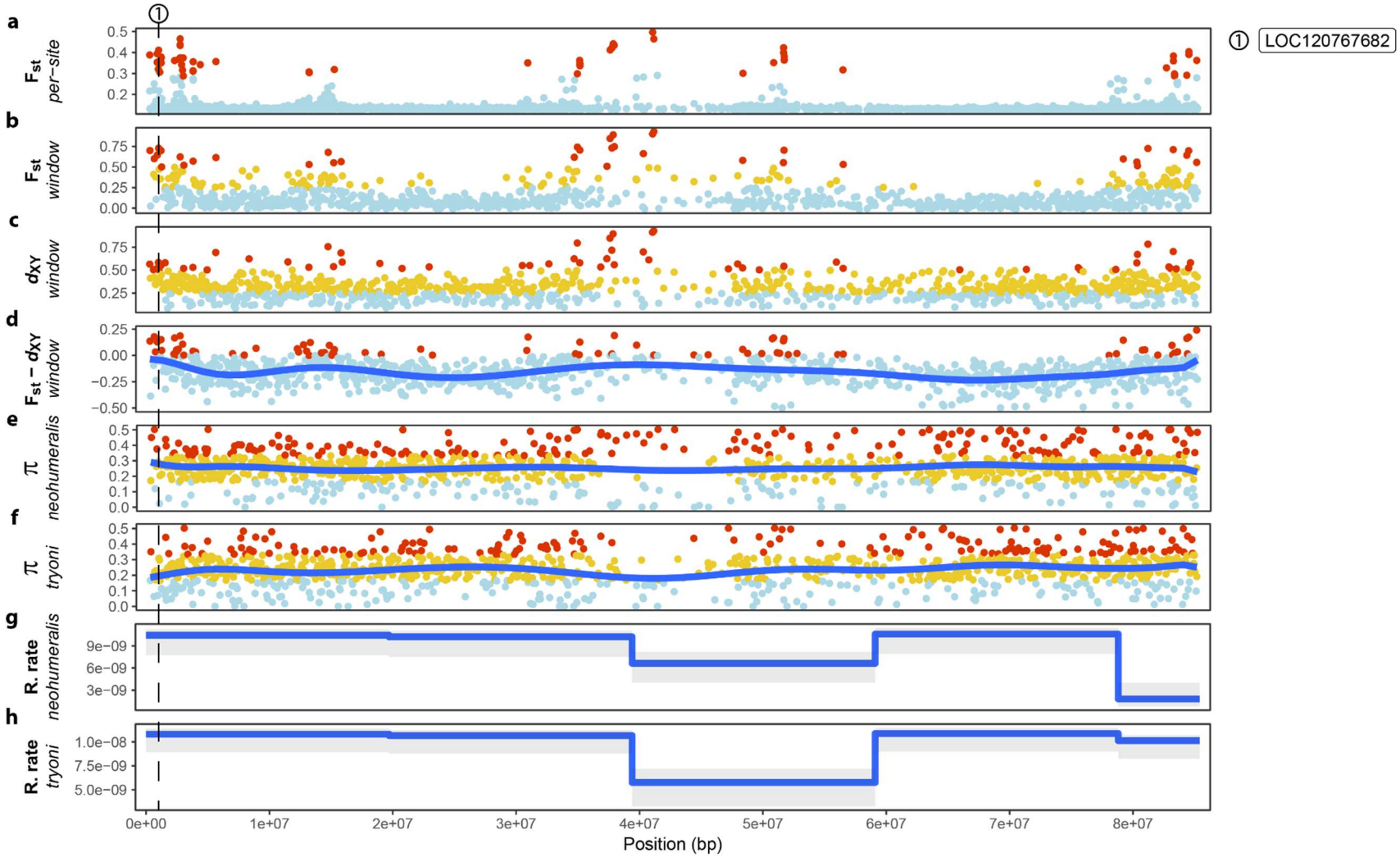


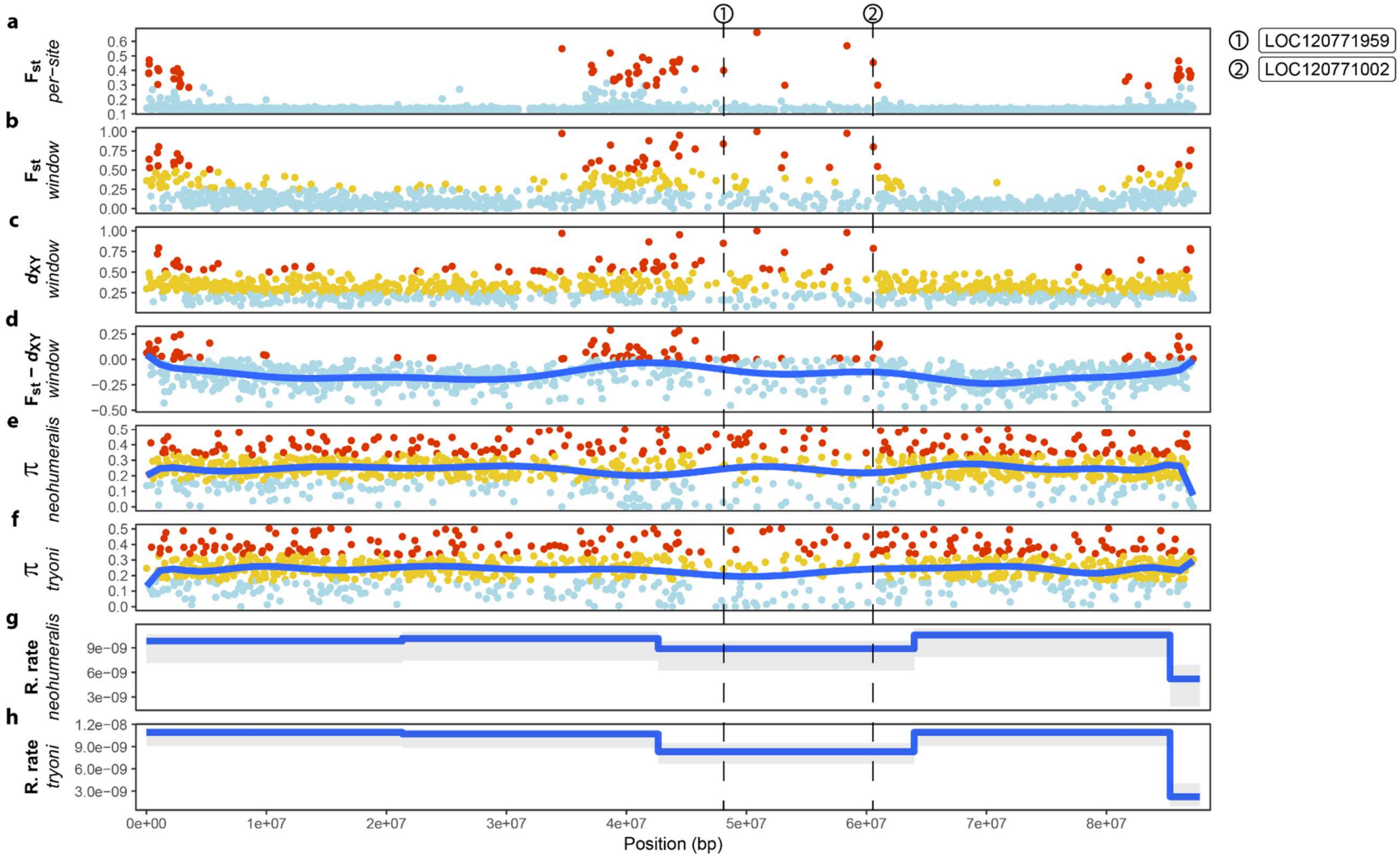
Figure 7. Population genetics statistics calculated for chromosome 1 labelled as NC_052499.1. Where indicated on the Y-axis that a statistic is windowed, values have been obtained through averaging over window lengths of 50 Kbp. Each statistic is presented along the

length of the chromosome where the position labelled as 1e+07 equates to 10 Mbp. Candidate genes potentially involved in the assortative mating of the two species are indicated as gapped vertical lines denoted by a number which relates to the locus identifiers shown in the upper right corner of the figure (See Table S4 for candidate gene details). For subfigures b and c, colours are depicted according to values being in the range of: 0 to 0.25 is light blue; >0.25 to 0.5 is yellow; >0.5 to 1 is red. For subfigures e and f, colours are depicted according to values being in the range of: 0 to 0.2 is light blue; >0.2 to 0.35 is yellow; >0.35 to 1 is red. a) F_{ST} calculated for each variant is indicated, with outlier variants coloured in red and non-outliers as light blue. b) F_{ST} calculated in windows inclusive of invariant sites. c) d_{XY} is calculated in windows inclusive of invariant sites. d) The values shown in subfigures b and c are subtracted to depict the difference between F_{ST} and d_{XY} values in each windowed region with the trendline in blue. e) Nucleotide diversity (π) is calculated in windows inclusive of invariant sites for *Bactrocera neohumeralis* with the trendline in blue. f) As above but showing *B. tryoni's* nucleotide diversity. g) The recombination rate as predicted by ReLERNN for *B. neohumeralis* is indicated with grey shading showing the 95% confidence interval. h) As above but showing *B. tryoni's* recombination rate.



435

436 Figure 8. Population genetics statistics calculated for chromosome 2 labelled as NC_052500.1. Figure details are described in Fig. 7.


a
Fst per-site
b
Fst window
c
dXY window
d
Fst − dXY window
e
π neohumeralis
f
π tryoni
g
R. rate neohumeralis
h
R. rate tryoni
① LOC120771959
② LOC120771002
Position (bp)

437

Figure 9. Population genetics statistics calculated for chromosome 3 labelled as NC_052501.1. Figure details are described in Fig. 7.

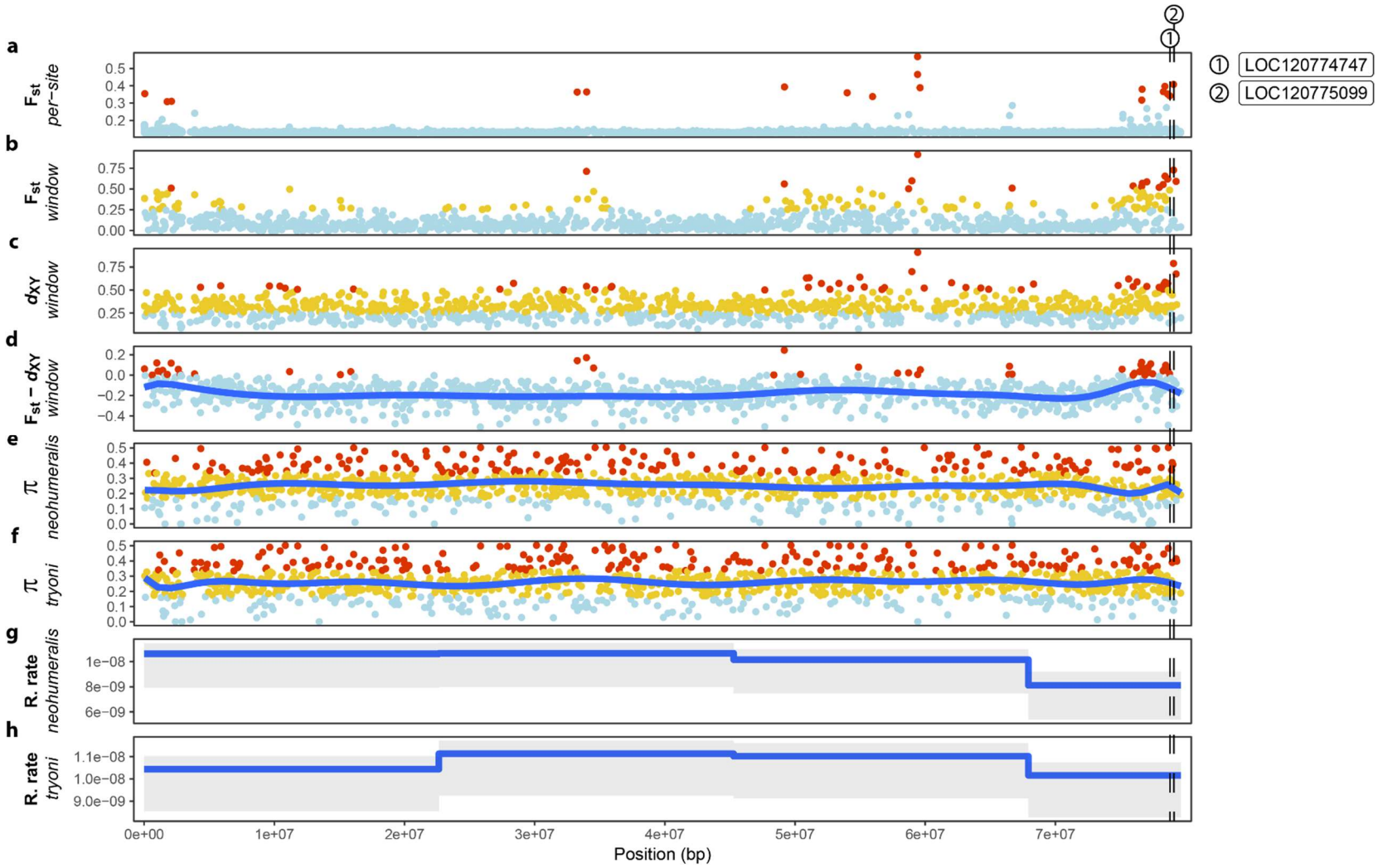
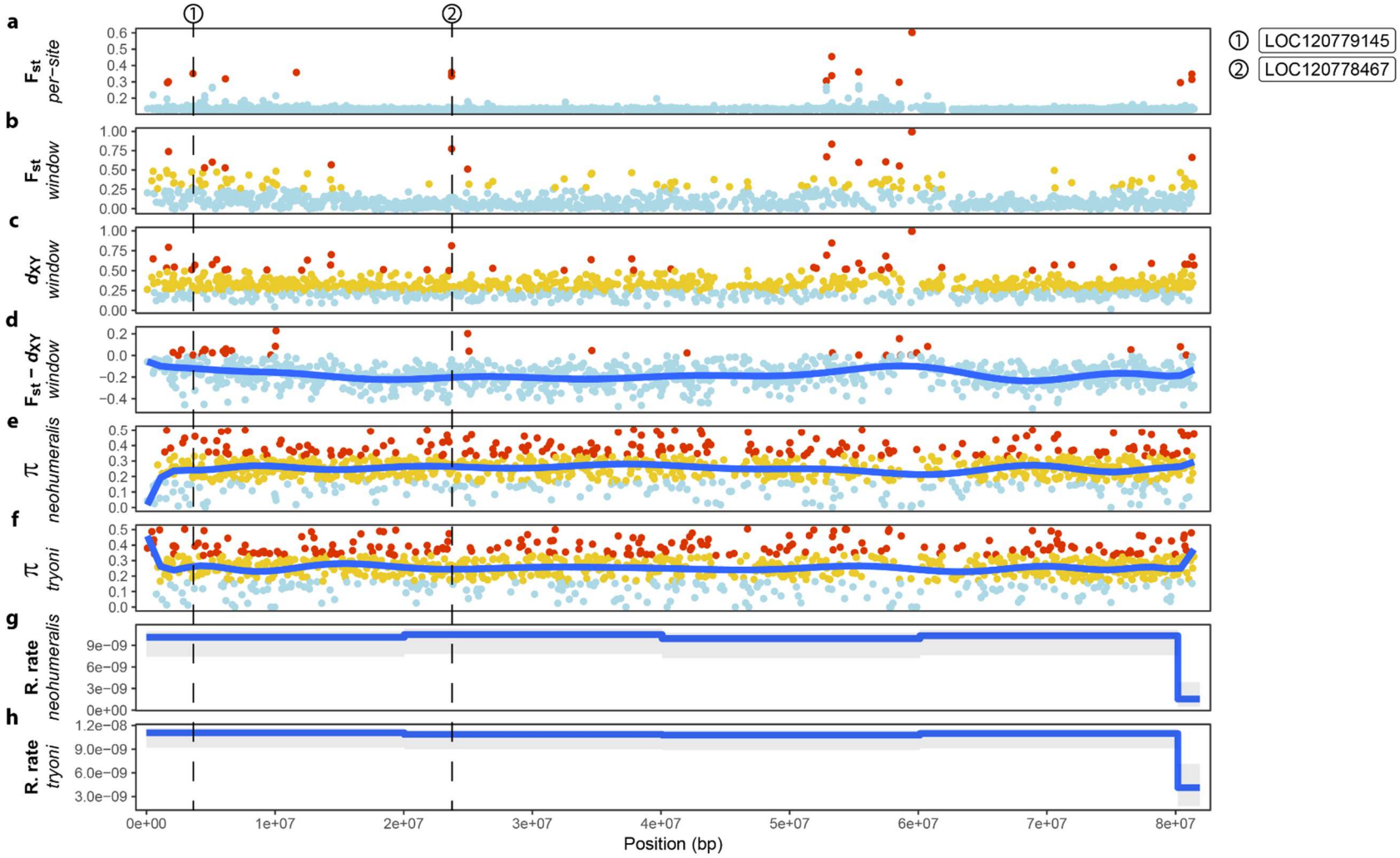


Figure 10. Population genetics statistics calculated for chromosome 3 labelled as NC_052502.1. Figure details are described in Fig. 7.



441

442 Figure 11. Population genetics statistics calculated for chromosome 3 labelled as NC_052503.1. Figure details are described in Fig. 7.

443 Mean nucleotide diversity estimates (π) produced similar results for each species (*B. tryoni* =
444 0.243962, *B. neohumeralis* = 0.242033). Nucleotide diversity estimates for each species were
445 homogeneously distributed along the chromosomes with no significant clustering of elevated
446 or lower rates observed (See Fig. 6 to Fig. 10). The LOSTRUCT results support the F_{ST} and
447 d_{XY} values with significant separation between species along PC1.

448 Overall, recombination rates within each species were low, with *B. tryoni* having a mean
449 value of 9.02×10^{-9} per base and *B. neohumeralis* having a mean value of 9.78×10^{-9} per base.
450 Recombination rates were overall consistently distributed along each chromosome with minor
451 decreases observed towards the telomeres and the centromere with eight regions of decreased
452 recombination rates occurring for each species (See Fig. 7 to Fig. 11). In each instance, these
453 regions included clustered regions of elevated F_{ST} and d_{XY} values, however not all clustered
454 regions of elevated F_{ST} and d_{XY} occurred in regions with lower recombination rates.

455 *Outlier loci are clustered in specific chromosome segments between the species*

456 Genes that contained outlier SNPs or were in close proximity to outlier SNPs were
457 investigated, resulting in the identification of 266 genes. Enrichment of annotation terms
458 indicates an overabundance of genes with the 'abnormal neuroanatomy' term (n=11 genes, p=
459 0.0164) within these candidate genes. Several neurology-related terms that were not enriched
460 but still of interest included 'nervous system development', 'central nervous system
461 development', 'sensory system', and 'sensory organ development'. Additional non-enriched
462 terms of interest included sleep and circadian rhythm-related terms, including: 'abnormal
463 sleep', 'circadian rhythm', 'abnormal circadian rhythm', and 'abnormal circadian behaviour'.
464 Inositol-related terms were statistically enriched, that included 'inositol biosynthetic process'

465 (n=1 gene, p= 0.0115) and ‘inositol-1,4,5-trisphosphate 3-kinase activity’ (n=1 gene, p=
466 0.0467). The 19 genes annotated with these terms are shown in Table S4.

467 Ecological Data

468 *No differences in seasonal activity between the species*

469 Broad scale trap data collected over multiple years and sites throughout the sympatric range
470 showed no difference in seasonal activity between *B. neohumeralis* and *B. tryoni* (Fig. 12 &
471 Fig. S4). For any given sampling event there were often significantly more *B. tryoni* than *B.*
472 *neohumeralis* trapped (See Table S5), but the abundance of both species over time were
473 highly linearly correlated across nine sites with correlation values ranging from: $r = 0.776$
474 (Ayr, $p < 0.001$) to $r = 0.993$ (Maryborough, $p < 0.001$) (See Table S6). South Johnstone was
475 the only site with a non-significant result ($r = 0.926$; $p = 0.074$) however, limited collections
476 were recorded at this site ($n = 4$), providing insufficient statistical power to detect an
477 association.

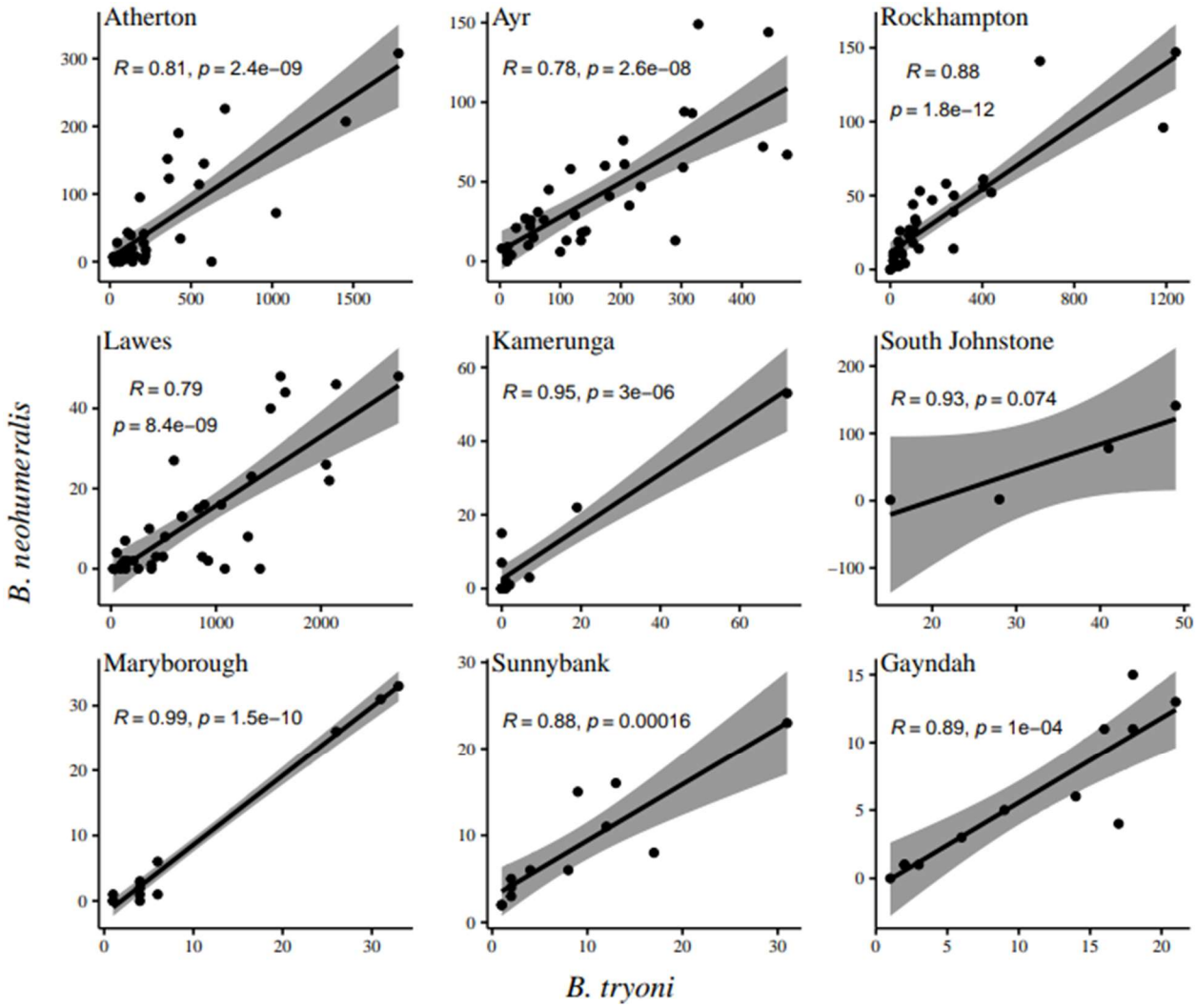


Figure 12. Correlation of seasonal abundance of *Bactrocera tryoni* and *B. neohumeralis* at multiple locations in Queensland, Australia. Scatterplots were developed using monthly mean of trap catches over 2 - 7 years. Shaded areas represent the 95% confidence intervals. Data from May (1961).

Trap data collected on a fine scale throughout the Wide Bay-Burnett region over a 12-month period shows that the abundance of both species across sites were significantly correlated over most collection periods with time, with the exception of the December period (r = 0.12, p = 0.57) where *B. neohumeralis* abundance decreased at a greater rate than *B. tryoni* after the

October collection (Fig. 13 & Table S7). As observed in the broad scale analyses, there were often more *B. tryoni* than *B. neohumeralis* collected (See Table S8).

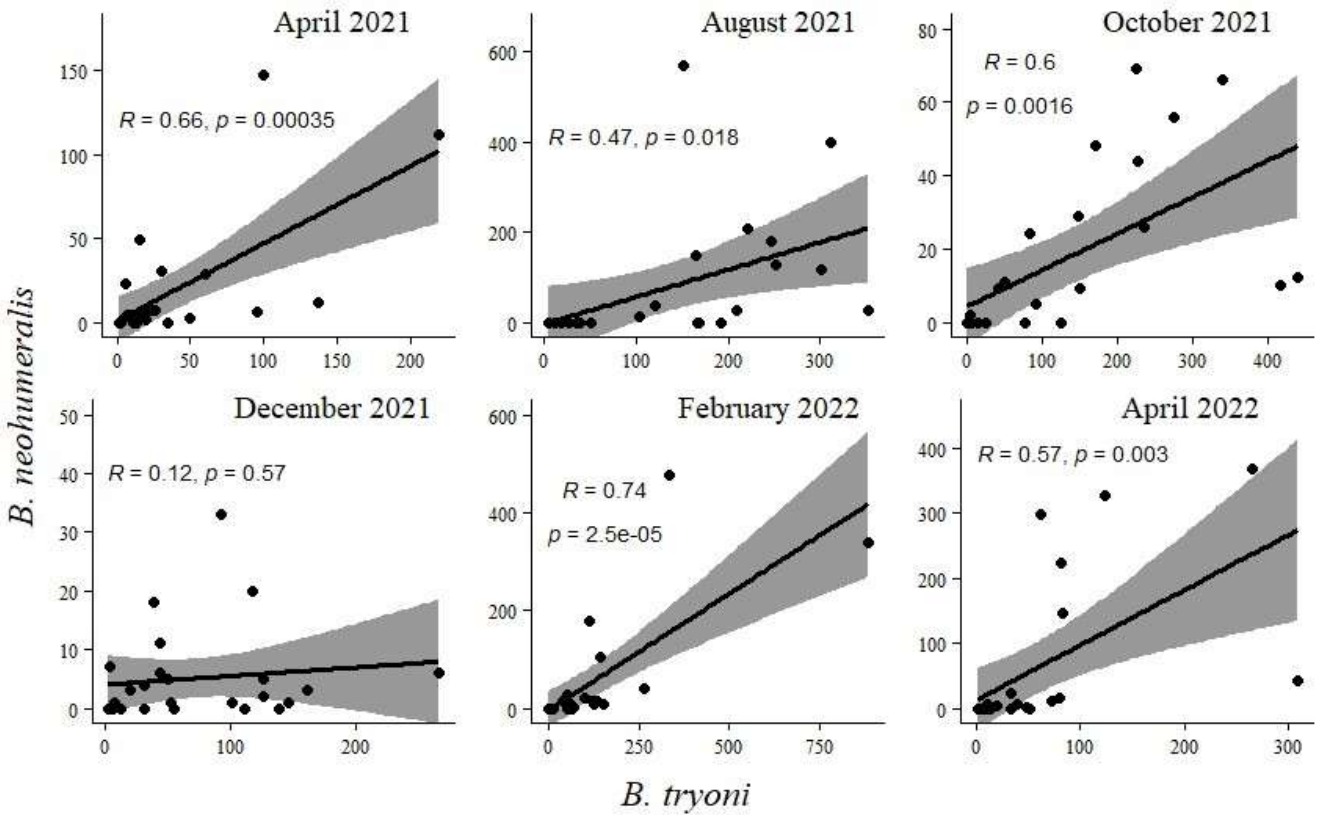


Figure 13: Correlation of fine scale seasonal abundance of *Bactrocera tryoni* and *B. neohumeralis* at multiple locations in the Wide Bay-Burnett region of Southeast Queensland. Shaded areas represent the 95% confidence intervals.

Correlated abundance across habitat types

There were significant positive correlations between *B. tryoni* and *B. neohumeralis* abundances in their habitat use (Fig. 14). Within a habitat type, there was a significant correlation in the two species monthly abundances for all habitat types, with the correlation ranging from r = 0.54 (dry forest, p <0.001) to r = 0.90 (mixed farming, p < 0.001) (See Table S10).

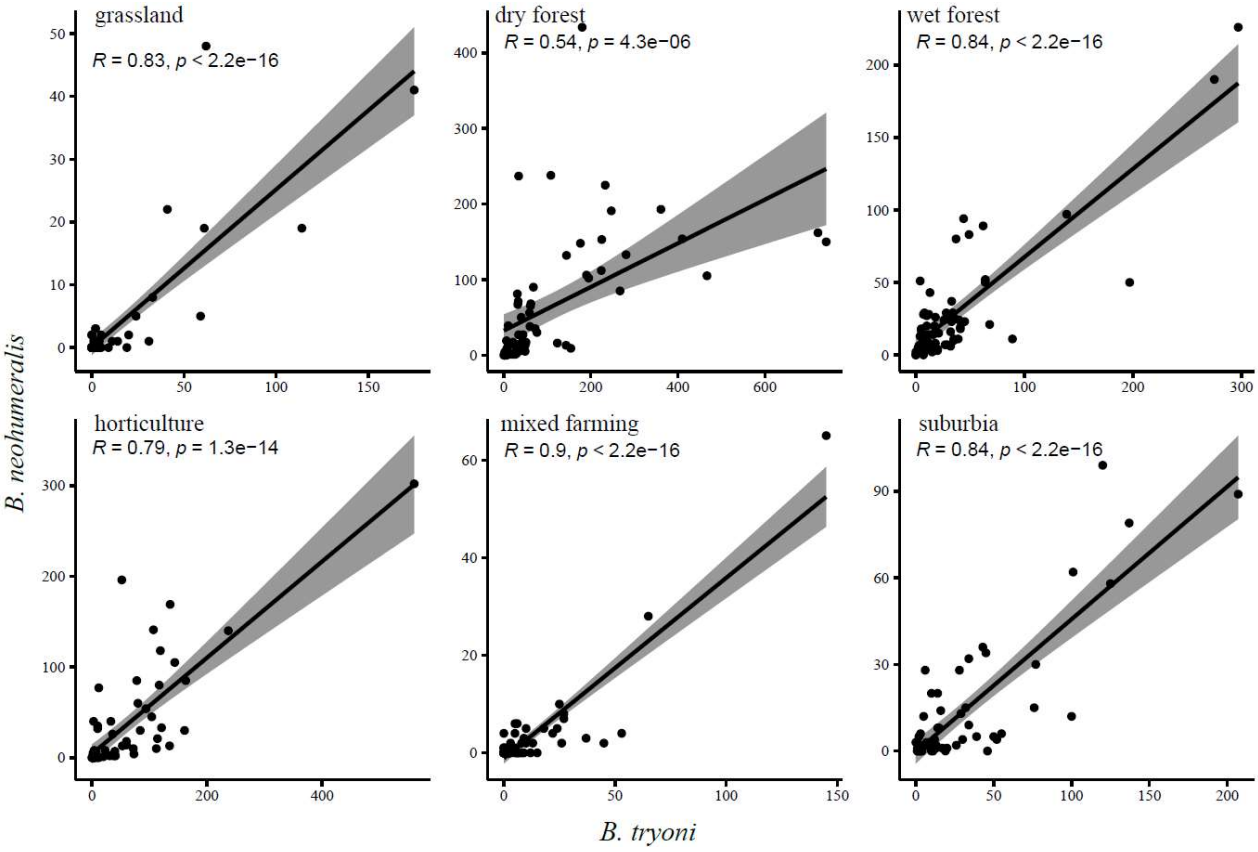


Figure 14. Correlation in habitat use of *Bactrocera tryoni* and *B. neohumeralis* in different human-defined habitat types within the Bundaberg district of Queensland. Each data point is the adult trap catch of both species for a trap run for three consecutive days. For each habitat type, nine traps were run per month for seven months. Shaded areas represent the 95% confidence intervals.

Landscape ecology

The LMM analysis included a total of 512 combinations of landscape predictors. Model averaging was performed using LMMs with ΔAICc < 2 from the top model (See Table S12). The top performing models retained eight predictors (See Table S13) from the nine included in the analysis, with only elevation not included in one of the top performing models. The top performing model was a combination of rainforest area, *Eucalyptus* woodland, latitude and

longitude, with each predictor positively correlated with *B. neohumeralis* abundance. The 13 other top performing LMMs all included rainforest area as a predictor. No other predictor was included in all the top performing LMMs. Following model averaging only rainforest area was a statistically significant predictor of *B. neohumeralis* abundance (See Table S13) and was positively correlated. This pattern can be observed in Figure 15 (Ai, ii, v, vi, vii), where the high levels of abundance observed correlate with rainforests. In contrast, Ryan et al., (in review) found grazing area was the only significant predictor of *B. tryoni* abundance and was negatively correlated (See Fig. 15 B.viii, Table S14, and Table S15).

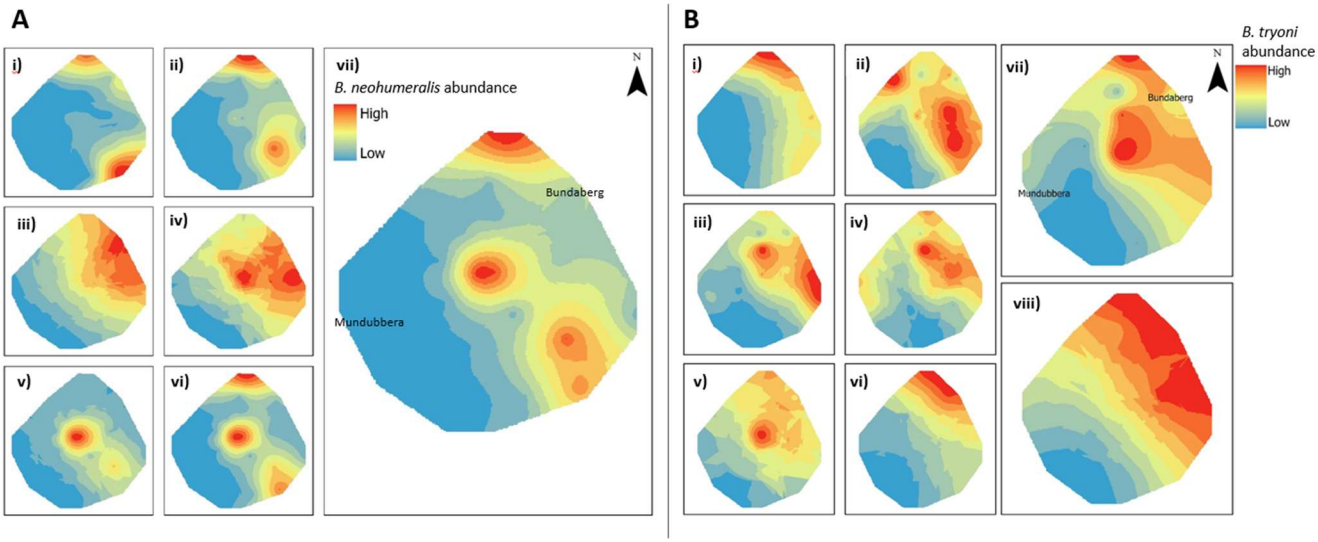
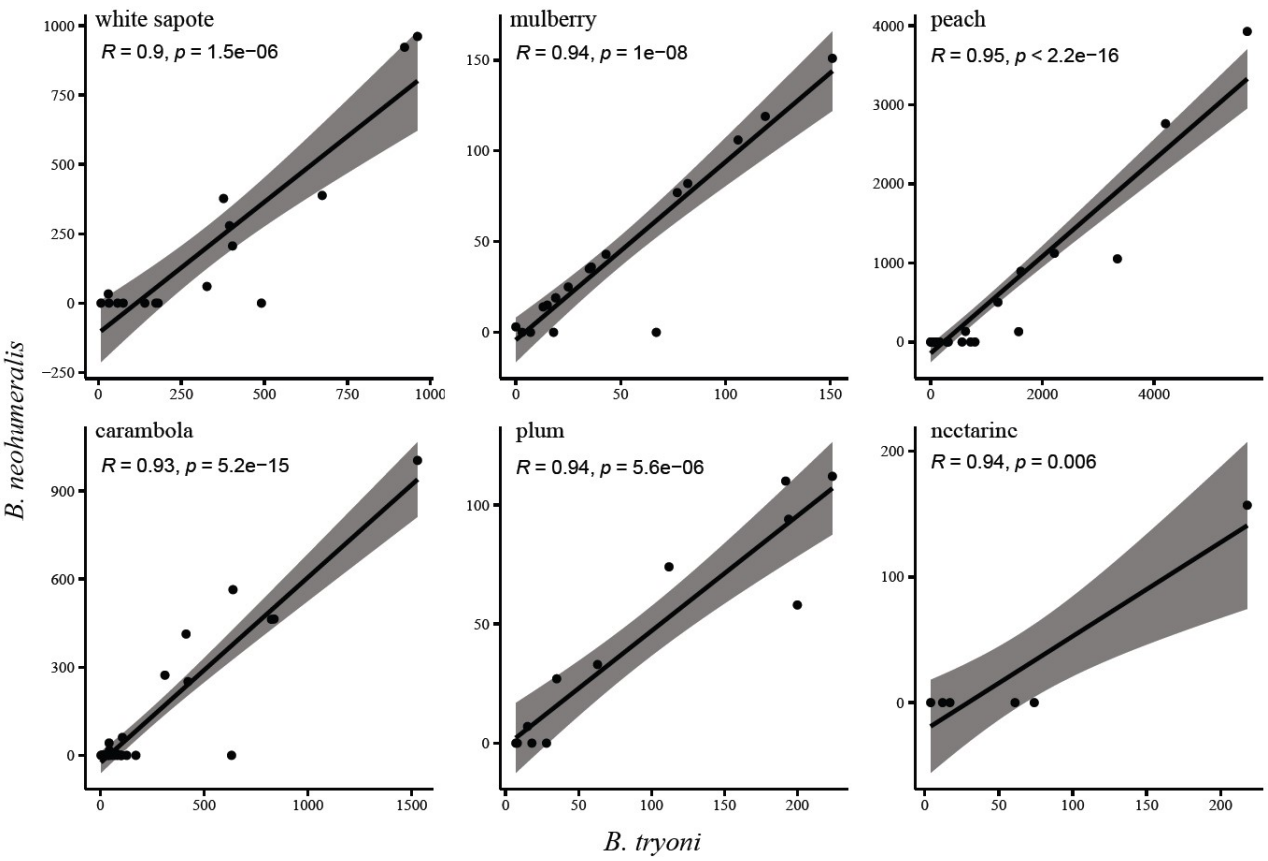


Figure 15: A) Empirical Bayesian kriging of *Bactrocera neohumeralis* abundance during each sampling period. i) April 2021, ii) August 2021, iii) October 2021, iv) December 2021, v) February 2022, vi) April 2022, vii) Average across all time periods. B) Empirical Bayesian kriging of *B. tryoni* abundance during each sampling period. i) April 2021, ii) August 2021, iii) October 2021, iv) December 2021, v) February 2022, vi) April 2022, vii) Average across all time periods, viii) Average abundance across all time periods excluding outlier from Good Night 2 in February 2022. Taken from Ryan et al. (In Review).

Limited differences in host use between species

A total of 1,607 fruits were sampled from the field and the emergence of *B. tryoni* and *B. neohumeralis* adults counted. For white sapote, mulberry, peach, plum, carambola and nectarine, sufficient samples were collected to carry out formal analyses. Highly significant correlations between the number of *B. tryoni* and *B. neohumeralis* adults reared from each of these fruit types were found (correlation ranging from r = 0.904, p <0.001 for white sapote to r = 0.952, p < 0.001 for peach) (Fig. 16 and Table S13). For guava, two collections occurred: a single *B. tryoni* and no *B. neohumeralis* emerged from the September 2013 collection while 1,609 *B. tryoni* and 1,353 *B. neohumeralis* emerged from the October 2013 collection; for feijoa a single collection occurred with equal numbers [71 flies each] of *B. tryoni* and *B. neohumeralis* emerging. For sapodilla, *Syzigium* spp., grumichama, hog plum, jabotica, white mulberry, avocado, black sapote, longan, mango and cashew, *B. tryoni* but no *B. neohumeralis* emerged.



542 Figure 16. Correlation in host use of *Bactrocera tryoni* and *B. neohumeralis*. Each data point
543 is the number of adult flies reared from an individual collection of multiple fruit pieces from a
544 research orchard at Nambour, Queensland. Shaded areas represent the 95% confidence
545 intervals.

546 **Discussion**

547 *High nucleotide diversity suggests allopatric divergence through vicariance*

548 Similar levels of nucleotide diversity (π) were estimated for *Bactrocera tryoni* (0.243962) and
549 *B. neohumeralis* (0.242033). These values are consistent with prior studies comparing the
550 nucleotide diversity of native (0.21552 – 0.22575) and invasive populations (0.19929 – 0.2199)
551 of *B. tryoni* throughout Australia and the South Pacific (Parvizi et al., 2024) but substantially
552 higher than those observed in other widespread *Bactrocera* species (*B. dorsalis* - 0.00148-
553 0.0183) (Deschepper et al., 2023). High effective population sizes and high genetic diversity have
554 previously been observed in population genetic studies on *B. tryoni* (Popa-Báez et al., 2020; Ryan
555 et al., In Review) which may account for the elevated levels of nucleotide diversity observed in
556 both species presented here. Low chromosome-wide recombination rates may also explain high
557 nucleotide diversity as genetic mutations, typically purged through recombination (Kondrashov,
558 1994), are able to accumulate. Furthermore, low recombination rates may allow for the
559 preservation of associations between linked alleles (Ortiz-Barrientos et al., 2016; Otto, 2009).
560 Yeap et al. (2020) observed a persistence in linkage between two *B. tryoni* - *B. neohumeralis*
561 hybrid phenotypes involving humeral calli colour and time-of-day mating over 25 generations
562 of hybrid backcrossing (without deliberate selection) despite a predicted linkage decay of
563 98.3% over 24 generations. The authors postulated that a reduction in recombination rates as a
564 result of chromosomal rearrangements likely contributed to the persistence of linkage, they also

565 suggested that an accumulation of nucleotide differences may also have been a factor. Low
566 recombination rates have been observed in a number of Diptera (Cladera, 1981; Foster et al.,
567 1980; Rössler, 1982) which can be attributed to an absence (or scarcity) of recombination in
568 males (Zhao et al., 2003).

569 Consistently high levels of nucleotide diversity along each of the chromosomes coupled with
570 predominantly low F_{ST} and d_{XY} values could be an indicator as to how initial divergence
571 between these species occurred. Under a scenario of speciation with gene flow or allopatric
572 speciation involving a small founder population, a genetic bottleneck is expected to reduce the
573 nucleotide diversity of the population expanding into the novel environment (Coyne & Orr, 2004).
574 The low recombination rates observed in *B. tryoni* and *B. neohumeralis* may have allowed for
575 the accumulation of high nucleotide diversity within each species post speciation, however, one
576 would expect to see elevated F_{ST} and d_{XY} values across the chromosome as a result, which is
577 not evident in our data. An alternative explanation is that a substantial vicariance event
578 separated the ancestral species into two large populations, allowing speciation to occur in the
579 absence of gene flow. During the period of geographic isolation, differential selection pressure
580 would have led to the differences in the time-of-day mating between the species, limiting
581 hybridization upon secondary contact (Coyne & Orr, 2004; Mayr, 1957).

582 *Genomic landscape indicates the timing of genetic divergence during speciation*

583 A number of tightly clustered regions of elevated genetic differentiation appear to occur in
584 regions with further reduced recombination along chromosomes 2-5 (See Fig. 8 to Fig. 11).
585 These regions are consistent with models of hybridizing species where the linkage of genes
586 involved in assortative mating with those associated with habitat preference and performance
587 is considered necessary for speciation with gene flow (Felsenstein, 1981; Ortíz-Barrientos et al.,
588 2002; Yeaman, 2013). The formation of the tightly clustered regions of genetic differentiation

589 observed in data presented here may have been associated with the initial divergence of the two
590 species if speciation occurred in sympatry. However, based on the high levels of nucleotide
591 diversity observed it is perhaps more likely that these tightly clustered regions were important
592 in maintaining species boundaries during the early periods of secondary contact where the
593 species diverged in allopatry. If reproductive barriers were porous during secondary contact,
594 the strong linkage of locally adapted alleles and genes associated with assortative mating would
595 have likely strengthened the reproductive barriers between the two species, limiting the
596 homogenizing effect of hybridization (Felsenstein, 1981; Ortíz-Barrientos et al., 2002).

597 In addition to those clustered regions with reduced rates of recombination, a number of regions
598 of elevated genetic diversity without a reduction in recombination rates are present. Elevated
599 genetic diversity without reduced rates of recombination are consistent with the patterns
600 expected under the model of divergence after speciation where populations experience a period
601 of geographic isolation allowing divergence to occur in the absence of gene flow (Cruickshank
602 & Hahn, 2014). Under this model, differential selective pressures drive divergence between the
603 species through selective sweeps and the absence of gene flow allows them to persist without a
604 reduction in recombination. The level of absolute genetic divergence (d_{XY}) in relation to relative
605 genetic differentiation (F_{ST}) could provide insight into the timing of these selective sweeps
606 relative to the completion of speciation between taxa. Where both statistics are elevated, the
607 selective sweep likely occurred soon after speciation. In contrast, regions with elevated F_{ST} and
608 non-elevated d_{XY} could suggest a more recent selective sweep (Cruickshank & Hahn, 2014; Ravinet
609 et al., 2017).

610 *Assortative mating maintains species boundaries in sympatry*

611 Genetic comparisons of *B. tryoni* and *B. neohumeralis* individuals collected throughout the
612 overlapping range showed no evidence of hybridization between the species. This was despite

samples of each species regularly being collected from the same traps, i.e., in immediate sympatry. Several intermediate phenotypes were collected, however, these samples presented strongly as *B. neohumeralis* genotypes with no indication of hybridization. This is consistent with the findings of prior studies of the species pair and supports the theory that strong reproductive barriers exist in natural conditions (Gilchrist & Ling, 2006; Pike, 2004; Yeap et al., 2020). The lack of evidence of hybridization identified from natural populations in this study is consistent with the results of a recent diagnostic tool development study where no hybrids were identified from *B. tryoni* (n = 430) and *B. neohumeralis* (n = 99) samples collected across > 40 collection sites throughout their overlapping range (Starkie et al., 2023). These results along with asymmetric hybridization observed in a prior laboratory-based study (Yeap et al., 2020), suggest that assortative mating is the most prevalent form of reproductive isolation responsible for maintaining species boundaries between the pair. However, the existence of hybrid selection in natural environments cannot be discounted, definitively.

While genes related to mating behaviour and circadian rhythm were not enriched in the gene enrichment analysis, 19 candidate genes were located in the tightly clustered regions of high genetic differentiation. These genes are postulated as being responsible for maintaining separation between the two species and forming the time-of-day premating reproductive barrier through assortative mating. Further analysis is required to determine whether these putative genes associated with assortative mating are linked to locally adapted genes, given the apparent lack of ecological niche differentiation between the species as discussed below.

Lack of competitive displacement allows co-existence

Competitive displacement has long been suggested as an important mechanism facilitating divergence and speciation (Rice & Salt, 1988; Thoday & Gibson, 1962; Via, 2001). The ability of two nascent species to co-exist by minimizing competition, or by utilizing non-competitive

637 ecological space, should be considered central to early species divergence and persistence
638 (Germain et al., 2021). Where two herbivore populations share resources, then displacement can
639 occur in time (i.e. seasonal activity), location, or host usage (Clarke & Measham, 2022; Facon et
640 al., 2021). However, the data presented here show no evidence for this in the *B. tryoni*/*B.*
641 *neohumeralis* sibling pair. Both species were active at the same time of the year and in the same
642 habitats within a landscape. *Bactrocera tryoni* was recovered from a greater range of host fruits
643 than *B. neohumeralis* and that may be seen as evidence for competitive displacement (i.e.
644 competitive advantage by *B. neohumeralis* resulting in *B. tryoni* adapting to novel hosts), but it
645 is more likely that this pattern simply reflects *B. tryoni*'s greater abundance in the environment
646 and so greater likelihood of being sampled, rather than a competition avoidance mechanism.
647 This is because when both species were reared from the same fruit, their numbers were closely
648 correlated. In cases where competition between fruit flies has been known to cause changes in
649 field host use patterns, one species will dominate in a given host, pushing a second species to
650 host fruits that the first only irregularly uses (Charlery de la Masselière et al., 2017; Hassani et al.,
651 2022); this pattern was not observed in these data. The positive correlation in the number of
652 adults of both species from fruit also offers field data supporting the work of Kay et al. (2024),
653 who in laboratory studies found no interspecific larval competition between the species.

654 Ecological traits typically associated with competition driven divergence between fruit fly
655 species (i.e. differential host use, habitat usage, and seasonal abundance (Clarke & Measham,
656 2022; Duyck et al., 2004; Inskeep et al., 2021)) do not appear to differ between taxa in this study,
657 however, the results of our LMM and those of Ryan et al. (In Review) indicate there may be
658 some minor difference in habitat preference. We show that *B. neohumeralis* appears to prefer
659 rainforest systems which is a subset of the habitat preference of *B. tyroni* which is only affected

660 by the presence of cleared grazing area. Therefore, the difference in mating time remains the
661 only known distinction between the species pair.

662 **Author contributions**

663 A.R.C. and P.P. conceived the study. M.I., N.K., C.G.M., J.R., S.B. and, B.M. conducted field
664 work and collected samples for molecular and ecological analysis. M.I., N.K., J.R., A.R.C.
665 and, P.P. designed, collated, and analyzed ecological data. Z.S. and P.P. designed and
666 performed bioinformatics and SNP analysis. M.I., Z.S., P.P. collated analyzed molecular data.
667 M.I., Z.S., N.K., J.R., C.G.M., M.S., A.R.C., D.H. and P.P. contributed to the writing of the
668 manuscript. All authors have read and approved the final version.

669 **Acknowledgements**

670 We thank E. Vueti who carried out much of the Bundaberg sampling as part of an
671 uncompleted postgraduate study, and L. Senior (QDAF) for her assistance with rearing flies
672 from fruit that provided data for the host utilization study. We also thank The Southern
673 Downs Regional Council, The Granite Belt Growers Association, B. Doohan, The
674 Greenhoose, Lockhart River, D. Masters, J. Kerr and J. Whistler for their assistance in
675 collecting samples for the molecular analysis.

676 This research was partially funded by the ARC ITTC project IC150100026, Centre for Fruit
677 Fly Biosecurity Innovation awarded to A.R.C. and P.P. During the final stages of writing
678 P.J.P., D.H., M.S., B.M., and S.B. received support through Fresh and Secure Trade Alliance.
679 The Fresh and Secure Trade Alliance is funded through the Hort Frontiers International
680 Markets Fund, part of the Hort Frontiers strategic partnership initiative developed by Hort
681 Innovation, with co-investment from the Department of Agriculture and Fisheries,
682 Department of Primary Industries and Regional Development, Department of Energy,

683 Environment and Climate Action, Department of Tourism, Industry and Trade, Department of
684 Primary Industries and Regions, Department of Natural Resources and Environment,
685 Queensland University of Technology, James Cook University, Western Sydney University,
686 Australian Blueberry Growers' Association, GreenSkinAvocados, and contributions from the
687 Australian Government and the strawberry and avocado R&D levy.

688 The authors declare no conflicts of interest.

689 **Information on Data Deposition**

690 BioProject PRJNA1008389, Zachary Stewart, 2023, DArTseq of *Bactrocera neohumeralis*
691 and *tryoni* (Queensland fruit fly / Qfly), NCBI SRA, BioProject PRJNA1008389.

692 **References**

693 Adrion, J. R., Galloway, J. G., & Kern, A. D. (2020). Predicting the landscape of
694 recombination using deep learning. *Molecular Biology and Evolution*, *37*(6), 1790-
695 1808. https://doi.org/10.1093/molbev/msaa038
696 Barton, K. (2022). *MuMIn: Multi-model inference* In R package version 1.47.1. https://cran.r-
697 project.org/package=MuMIn
698 Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects
699 models using lme4. *Journal of statistical software*, *67*(1), 1-48.
700 https://doi.org/10.18637/jss.v067.i01
701 Bellas, T., & Fletcher, B. S. (1979). Identification of the major components in the secretion
702 from the rectal pheromone glands of the Queensland fruit flies, *Dacus tryoni* and
703 *Dacus neohumeralis* (Diptera: Tephritidae). *Journal of Chemical Ecology*, *5*(5), 795-
704 803.
705 Birch, L. C. (1961). Natural selection between two species of tephritid fruit flies of the genus
706 *Dacus*. *Evolution*, *15*, 360-374.
707 Birch, L. C., & Vogt, W. G. (1970). Plasticity of taxonomic characters of the Queensland fruit
708 flies *Dacus tryoni* and *Dacus neohumeralis* (Tephritidae). *Evolution*, *24*, 320-343.

Butlin, R. K. (2005). Recombination and speciation. *Molecular Ecology*, *14*(9), 2621-2635. https://doi.org/10.1111/j.1365-294X.2005.02617.x

Charlery de la Masselière, M., Ravigné, V., Facon, B., Lefeuvre, P., Massol, F., Quilici, S., & Duyck, P.-F. (2017). Changes in phytophagous insect host ranges following the invasion of their community: Long-term data for fruit flies. *Ecology and Evolution*, *7*(14), 5181-5190. https://doi.org/https://doi.org/10.1002/ece3.2968

Cladera, J. L. (1981). Absence of recombination in the male of *Ceratitis capitata*. *Experientia*, *37*(4), 342-343. https://doi.org/10.1007/BF01959852

Clarke, A. R., Leach, P., & Measham, P. F. (2022). The fallacy of year-round breeding in polyphagous tropical fruit flies (Diptera: Tephritidae): Evidence for a seasonal reproductive arrestment in *Bactrocera* species. *Insects*, *13*, NA.

Clarke, A. R., & Measham, P. F. (2022). Competition: A missing component of fruit fly (Diptera: Tephritidae) risk assessment and planning. *Insects*, *13*.

Cowley, J. M., Page, F. D., Nimmo, P. R., & Cowley, D. R. (1990). Comparison of the effectiveness of two traps for *Bactrocera tryoni* (Froggatt)(Diptera:Tephritidae) and implications for quarantine surveillance systems. *Australian Journal of Entomology*, *29*(3), 171-176. https://doi.org/https://doi.org/10.1111/j.1440-6055.1990.tb00342.x

Coyne, J. A., & Orr, H. A. (2004). *Speciation* (Vol. 37). Sinauer Associates Sunderland, MA.

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133-3157. https://doi.org/https://doi.org/10.1111/mec.12796

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

Deschepper, P., Vanbergen, S., Zhang, Y., Li, Z., Hassani, I. M., Patel, N. A., Rasolofoarivao, H., Singh, S., Wee, S. L., De Meyer, M., Virgilio, M., & Delatte, H. (2023).

Bactrocera dorsalis in the Indian Ocean: A tale of two invasions. *Evolutionary applications*, *16*(1), 48-61. https://doi.org/https://doi.org/10.1111/eva.13507

Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, *21* 113-135.

Drew, R., Hooper, G., & Bateman, M. (1982). *Economic fruit flies of the South Pacific Region*.

Drew, R. A. I. (1989). The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and Oceanian regions. In *Memoirs of the Queensland Museum* (Vol. 26, pp. 1-521).

Duyck, P.-F., David, P., & Quilici, S. (2004). A review of relationships between interspecific competition and invasions in fruit flies (Diptera: Tephritidae). *Ecological Entomology*, *29*(5), 511-520. https://doi.org/https://doi.org/10.1111/j.0307-6946.2004.00638.x

Ekanayake, W. M. T. D., Jayasundara, M. S. H., Peek, T., Clarke, A. R., & Schutze, M. K. (2017). The mating system of the true fruit fly *Bactrocera tryoni* and its sister species, *Bactrocera neohumeralis*. *Insect Science*, *24*(3), 478-490. https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/1744-7917.12337?download=true

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, *4*(3). https://doi.org/https://doi.org/10.3835/plantgenome2011.08.0024

Facon, B., Hafsi, A., Charlery de la Masselière, M., Robin, S., Massol, F., Dubart, M., Chiquet, J., Frago, E., Chiroleu, F., Duyck, P.-F., & Ravigné, V. (2021). Joint species distributions reveal the combined effects of host plants, abiotic factors and species competition as drivers of species abundances in fruit flies. *Ecology Letters*, *24*(9), 1905-1916. https://doi.org/https://doi.org/10.1111/ele.13825

Feder, J. L., Flaxman, S. M., Egan, S. P., Comeault, A. A., & Nosil, P. (2013). Geographic mode of speciation and genomic divergence. *Annual Review of Ecology, Evolution, and Systematics*, *44*(Volume 44, 2013), 73-97. https://doi.org/https://doi.org/10.1146/annurev-ecolsys-110512-135825

Felsenstein, J. (1981). Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, *35*(1), 124-138. https://doi.org/https://doi.org/10.1111/j.1558-5646.1981.tb04864.x

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both bominant and codominant markers: a bayesian perspective. *Genetics (Austin)*, *180*(2), 977-993. https://doi.org/10.1534/genetics.108.092221

Foster, G. G., Maddern, R. H., & Mills, A. T. (1980). Genetic instability in mass-rearing colonies of a sex-linked translocation strain of *Lucilia cuprina* (Wiedemann) (Diptera: Calliphoridae) during a field trial of genetic control. *Theoretical and Applied Genetics*, *58*(3), 169-175. https://doi.org/10.1007/BF00263113

Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27-32. https://doi.org/10.1111/1755-0998.12509

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. https://doi.org/10.48550/arxiv.1207.3907

Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: contrasting theory with data. *Science*, *323*(5915), 732-737. https://doi.org/10.1126/science.1157966

Germain, R. M., Hart, S. P., Turcotte, M. M., Otto, S. P., Sakarchi, J., Rolland, J., Usui, T., Angert, A. L., Schluter, D., Bassar, D. D., Waters, M. T., Henao-Diaz, F., & Siepielski, A. M. (2021). On the origin of coexisting species. *Trends in Ecology & Evolution*, *36*, 284-293.

Gilchrist, A. S., Cameron, E. C., An, X., Meats, A., Frommer, M., & Raphael, K. A. (2005). The tryoni complex of tephritid flies in Australia. In *Proceedings of 2nd FAO/IAEA International Conference on Area-Wide Control of Insect Pests: Integrating the Sterile Insect and Related Nuclear and Other Techniques,* . IAEA.

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*(2), 157-162. https://doi.org/10.1111/j.1440-6055.2006.00522.x

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*, Article 1153. https://doi.org/10.1186/1471-2164-15-1153

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome

804 sequencing. *Molecular Ecology Resources*, *18*(3), 691-699.
805 https://doi.org/10.1111/1755-0998.12745
806 Hancock, D. L., E.L., H., Lloyd, A. C., & Elson-Harris, M. M. (2000). *The distribution and*
807 *host plants of fruit flies (Diptera: Tephritidae) in Australia*. Queensland Department
808 of Primary Industries.
809 Hassani, I. M., Delatte, H., Ravaomanarivo, L. H. R., Nouhou, S., & Duyck, P.-F. (2022).
810 Niche partitioning via host plants and altitude among fruit flies following the invasion
811 of *Bactrocera dorsalis*. *Agricultural and Forest Entomology*, *24*(4), 575-585.
812 https://doi.org/https://doi.org/10.1111/afe.12522
813 Hey, J. (2001). *Genes, categories, and species : The evolutionary and cognitive cause of the*
814 *species problem*. Oxford University Press, Incorporated.
815 Inskeep, K. A., Doellman, M. M., Powell, T. H. Q., Berlocher, S. H., Seifert, N. R., Hood, G.
816 R., Ragland, G. J., Meyers, P. J., & Feder, J. L. (2021). Divergent diapause life history
817 timing drives both allochronic speciation and reticulate hybridization in an adaptive
818 radiation of Rhagoletis flies. *Molecular Ecology*, *n/a*(n/a).
819 https://doi.org/https://doi.org/10.1111/mec.15908
820 James, M. E., Arenas-Castro, H., Groh, J. S., Allen, S. L., Engelstädter, J., & Ortiz-
821 Barrientos, D. (2021). Highly replicated evolution of parapatric ecotypes. *Molecular*
822 *Biology and Evolution*, *38*(11), 4805-4821. https://doi.org/10.1093/molbev/msab207
823 Janz, N., & Nylin, S. (2008). The oscillation hypothesis of host-plant range and speciation. In
824 K. Tilmon (Ed.), *Specialization, speciation, and radiation: the evolutionary biology of*
825 *herbivorous insects*. University of California Press Berkeley.
826 Jeffries, M. J., & Lawton, J. H. (1984). Enemy free space and the structure of ecological
827 communities. *Biological journal of the Linnean Society*, *23*(4), 269-286.
828 https://doi.org/https://doi.org/10.1111/j.1095-8312.1984.tb00145.x
829 Kassambara, A. (2023). *ggpubr: 'ggplot2' based publication ready plots*. Retrieved 0 from
830 https://rpkgs.datanovia.com/ggpubr/
831 Kay, B., Harris, C., & Clarke, A. R. (2024). Larval competition between three endemic fruit
832 flies (Diptera: Tephritidae) of differing phylogenetic relatedness. *Austral Entomology:*
833 *in press.*
834 Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska,
835 K., Jaccoud, D., Hopper, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla,

C., Hok, P., & Uszynski, G. (2012). Diversity Arrays Technology: A generic genome profiling technology on open platforms. In F. Pompanon & A. Bonin (Eds.), *Data Production and Analysis in Population Genomics: Methods and Protocols* (pp. 67-89). Humana Press. https://doi.org/10.1007/978-1-61779-870-2_5

Kondrashov, A. S. (1994). Muller's ratchet under epistatic selection. *Genetics*, *136*(4), 1469-1473. https://doi.org/10.1093/genetics/136.4.1469

Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, *21*(4), 1359-1368. https://doi.org/https://doi.org/10.1111/1755-0998.13326

Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., dos Santos, G., Garapati, P. V., Goodman, Joshua L., Gramates, L S., Millburn, G., Strelets, V. B., Tabone, C. J., & Thurmond, J. (2021). FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic acids research*, *49*(D1), D899-D907. https://doi.org/10.1093/nar/gkaa1026

Leblanc, L., Vueti, E. T., Drew, R. A., & Allwood, A. J. (2012). Host plant records for fruit flies (Diptera: Tephritidae: Dacini) in the Pacific Islands. *Proceedings of the Hawaiian Entomological Society*, *44*, 11-53.

Lewontin, R. C., & Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution*, *20*(3), 315-336. https://doi.org/https://doi.org/10.1111/j.1558-5646.1966.tb03369.x

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*. https://doi.org/10.48550/arxiv.1303.3997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., & Ralph, P. (2018). Local PCA shows how the effect of population structure differs along the genome. *Genetics*, *211*(1), 289-304. https://doi.org/10.1534/genetics.118.301747

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298-299. https://doi.org/10.1093/bioinformatics/btr642

M'Gonigle, L. K., Mazzucco, R., Otto, S. P., & Dieckmann, U. (2012). Sexual selection enables long-term coexistence despite ecological equivalence. *Nature*, *484*(7395), 506-509. https://doi.org/10.1038/nature10971

Mankin, R. W., Lemon, M., Harmer, A. M. T., 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*(3), 664-674.

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.

Mayr, E. (1957). Species concepts and definitions. In E. Mayr (Ed.), *The Species Problem* (pp. 1-22). American Association for Advancement of Science.

Meats, A., Pike, N., An, X., Raphael, K., & Wang, W. Y. S. (2003). The effects of selection for early (day) and late (dusk) mating lines of hybrids of *Bactrocera tryoni* and *Bactrocera neohumeralis*. *Genetica*, *119*(3), 283-293. https://doi.org/10.1023/B:GENE.0000003683.42395.51

Mijangos, J. L., Gruber, B., Berry, O., Pacioni, C., & Georges, A. (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution*, *13*(10), 2150-2158. https://doi.org/10.1111/2041-210X.13918

Morrow, J. L., 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). Seasonal phenology of *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae) in Queensland. *Australian Journal of Entomology*, *49*(3), 221-233. https://doi.org/10.1111/j.1440-6055.2010.00759.x

Ortiz-Barrientos, D., Engelstädter, J., & Rieseberg, L. H. (2016). Recombination rate evolution and the origin of species. *Trends in Ecology & Evolution*, *31*(3), 226-236. https://doi.org/10.1016/j.tree.2015.12.016

Ortíz-Barrientos, D., Reiland, J., Hey, J., & Noor, M. A. (2002). Recombination and the divergence of hybridizing species. *Genetica*, *116*(2-3), 167-178. https://doi.org/https://doi.org/10.1023/A:1021296829109

Osborne, R., Meats, A., Frommer, M., Sved, J. A., Drew, R. A. I., & Robson, M. K. (1997). Australian distribution of 17 species of fruit flies (Diptera: Tephritidae) caught in cue lure traps in February 1994. *Australian Journal of Entomology*, *36*(1), 45-50.

Otto, Sarah P. (2009). The Evolutionary Enigma of Sex. *The American naturalist*, *174*(S1), S1-S14. https://doi.org/10.1086/599084

Parvizi, E., Vaughan, A. L., Dhami, M. K., & McGaughran, A. (2024). Genomic signals of local adaptation across climatically heterogenous habitats in an invasive tropical fruit fly (*Bactrocera tryoni*). *Heredity*, *132*(1), 18-29. https://doi.org/10.1038/s41437-023-00657-y

Pike, N. (2004). Natural incidence of fruit flies with character states intermediate to those of the sibling species *Bactrocera tryoni* (Froggatt) and *B. neohumeralis* (Hardy) (Diptera: Tephritidae). *Australian Journal of Entomology*, *43*(1), 23-27. https://doi.org/10.1111/j.1440-6055.2003.00377.x

Plant Health Australia. (2018). *The Australian handbook for the identification of Fruit Flies* (3.0 ed.). Plant Health Australia.

Plant Health Australia. (2023). *Fruit Fly ID Australia: Methods for detection*. https://www.fruitflyidentification.org.au/surveillance-and-detection/methods-for-detection/

Popa-Báez, Á.-D., Catullo, R., Lee, S. F., Yeap, H. L., Mourant, R. G., Frommer, M., Sved, J. A., Cameron, E. C., Edwards, O. R., Taylor, P. W., & Oakeshott, J. G. (2020). Genome-wide patterns of differentiation over space and time in the Queensland fruit fly. *Scientific Reports*, *10*(1), 10788. https://doi.org/10.1038/s41598-020-67397-5

QDPC. (2021). *Queensland Primary Industries Insect Collection Database.*

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics (Austin)*, *197*(2), 573-589. https://doi.org/10.1534/genetics.114.164350

Raphael, K. A., Sved, J. A., Pearce, S., Oakeshott, J. G., Gilchrist, A. S., Sherwin, W. B., & Fromme, M. (2019). Differences in gene regulation in a tephritid model of prezygotic reproductive isolation. *Insect molecular biology*, *28*, 689-702.

929 Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F.,
930 Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of
931 speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary*
932 *Biology*, *30*(8), 1450-1477. https://doi.org/10.1111/jeb.13047
933 Rice, W. R., & Salt, G. W. (1988). Speciation via disruptive selection on habitat preference:
934 experimental evidence. *The American naturalist*, *131*, 911-917.
935 Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods
936 for paired-end sequencing improve RADseq-based population genomics. *Molecular*
937 *Ecology*, *28*(21), 4737-4754.
938 Rössler, Y. (1982). Genetic recombination in males of the Mediterranean Fruit Fly, and Its
939 relation to automated sexing methods. *Annals of the Entomological Society of*
940 *America*, *75*(1), 28-31. https://doi.org/10.1093/aesa/75.1.28
941 Ryan, J., Clarke, A. R., Piper, A. M., Fuller, S., & Prentis, P. J. (In Review). Gene flow and
942 abundance of a tropical fruit fly in a horticultural landscape mosaic in eastern
943 Australia is limited by cleared grazing land and area-wide management.
944 Schluter, D. (2000). *The ecology of adaptive radiation*. OUP Oxford.
945 Schluter, D., & Rieseberg, L. H. (2022). Three problems in the genetics of speciation by
946 selection. *Proceedings of the National Academy of Sciences - PNAS*, *119*(30), 1-
947 e2122153119. https://doi.org/10.1073/pnas.2122153119
948 Smith, P. H. (1979). Genetic manipulation of the circadian clock's timing of sexual behaviour
949 in the Queensland fruit flies, *Dacus tryoni* and *Dacus neohumeralis*. *Physiological*
950 *Entomology*, *4*, 71-78.
951 Starkie, M. L., Fowler, E. V., Piper, A. M., Zhu, X., Wyatt, P., Gopurenko, D., Krosch, M. N.,
952 Strutt, F., Armstrong, K. F., Patrick, H., Schutze, M. K., & Blacket, M. J. (2023). A
953 novel diagnostic gene region for distinguishing between two pest fruit flies:
954 *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy) (Diptera:
955 Tephritidae). *Insect Science*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/1744-
956 7917.13299
957 Steiner, L. F. (1957). Low-cost plastic fruit fly trap. *Journal of Economic Entomology*, *50*(4),
958 508-509.
959 Tan, A., Abecasis, G. R., & Kang, H. M. (2015). Unified representation of genetic variants.
960 *Bioinformatics*, *31*(13), 2202-2204. https://doi.org/10.1093/bioinformatics/btv112

Thoday, J. M., & Gibson, J. B. (1962). Isolation by disruptive selection. *Nature*, *193*, 1164.

Turelli, M., Barton, N. H., & Coyne, J. A. (2001). Theory and speciation. *Trends in Ecology & Evolution*, *16*, 330-343.

Vayssières, J. F., De Meyer, M., Ouagoussounon, I., Sinzogan, A., Adandonon, A., Korie, S., Wargui, R., Anato, F., Houngbo, H., Didier, C., De Bon, H., & Goergen, G. (2015). Seasonal qbundance of Mango Fruit Flies (Diptera: Tephritidae) and ecological implications for their management in mango and cashew orchards in Benin (Centre & North). *Journal of Economic Entomology*, *108*(5), 2213-2230. https://doi.org/https://doi.org/10.1093/jee/tov143

Via, S. (2001). Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution*, *16*, 381-390.

Vogt, W. (1977). A re-evaluation of introgression between *Dacus Tryoni* and *Dacus Neohumeralis* (Diptera: Tephritidae). *Australian Journal of Zoology*, *25*(1), 59-69. https://doi.org/https://doi.org/10.1071/ZO9770059

Wang, J. (2022). A joint likelihood estimator of relatedness and allele frequencies from a small sample of individuals. *Methods in Ecology and Evolution*, *13*(11), 2443-2462. https://doi.org/https://doi.org/10.1111/2041-210X.13963

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*(4), 351-360.

Weber, M. G., & Strauss, S. Y. (2016). Coexistence in close relatives: Beyond competition and reproductive isolation in sister taxa. *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), 359-381. https://doi.org/10.1146/annurev-ecolsys-112414-054048

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358-1370. https://doi.org/10.2307/2408641

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.

Yeaman, S. (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences*, *110*(19), E1743-E1751. https://doi.org/doi:10.1073/pnas.1219381110

Yeap, H. L., Lee, S. F., Robinson, F., Mourant, R. G., Sved, J. A., Frommer, M., Papanicolaou, A., Edwards, O. R., & Oakeshott, J. G. (2020). Separating two tightly linked species-defining phenotypes in *Bactrocera* with hybrid recombinant analysis. *BMC Genetics*, *21*, 132. https://doi.org/10.1186/s12863-020-00936-1

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., Hagey, T. J., Jochimsen, D., Oswald, B. P., Robertson, J., Sarver, B. A. J., Schenk, J. J., Spear, S. F., & Harmon, L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, *23*(8), 1581-1596. https://doi.org/https://doi.org/10.1111/j.1420-9101.2010.02029.x

Zhao, J. T., Frommer, M., Sved, J. A., & Gillies, C. B. (2003). Genetic and molecular markers of the Queensland Fruit Fly, *Bactrocera tryoni*. *Journal of Heredity*, *94*(5), 416-420. https://doi.org/10.1093/jhered/esg088

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326-3328. https://doi.org/10.1093/bioinformatics/bts606