The redlegged earth mite draft genome provides new insights into pesticide resistance evolution and demography in its invasive Australian range

Joshua A. Thia [1,*], Pasi K. Korhonen [2], Neil D. Young [2], Robin B. Gasser [2], Paul A. Umina [3], Qiong Yang [1], Owain Edwards [4], Tom Walsh [5,6], Ary A. Hoffmann [1]

[1] Bio21 Institute, School of BioSciences, The University of Melbourne, Melbourne, Victoria 3052, Australia

[2] Department of Veterinary Biosciences, Melbourne Veterinary School, The University of Melbourne, Melbourne, Victoria 3052, Australia

[3] Cesar Australia, Brunswick, Victoria 3056, Australia

[4] CSIRO, Land and Water, Floreat, Western Australia 6014, Australia

[5] CSIRO, Black Mountain Laboratories, Canberra, Australian Capital Territory 2601, Australia

[6] Applied BioSciences, Macquarie University, Sydney, New South Wales 2109, Australia

[*] Corresponding author: joshua.thia@unimelb.edu.au, josh.thia@live.com

Running title: Genomics of invasive *Halotydeus destructor*

Keywords: acetylcholinesterase, Australia, copy number variation, demographic inference, *Halotydeus destructor*, invasive species, population genomics

Abstract

We present the first draft genome for an invasive mite, the redlegged earth mite, *Halotydeus destructor* (Tucker, 1925), a major pest of Australian grains and pasture. Our *de novo* assembly comprised 132 contigs, with a total length of 48.90 Mb, an N50 of 874,492, an L50 of 14, a GC content of 45.3%, 14,583 putative genes, and a high level of completeness as assessed through benchmark single copy orthologs (BUSCO >89%). We use our draft genome to gain insights into the unresolved genetic mechanism of organophosphate resistance and demographic processes in the invasive Australian populations. We found that the acetylcholinesterase gene (*ace*), the target of organophosphate pesticides, has undergone a large amplification event in *H. destructor*. Copy number of this gene is considerably larger compared to most other arthropod pests and is likely to contribute to organophosphate resistance in this species. Congruent with a stepping-stone invasion pathway, we found that western populations were more genetically diverse (higher genomic heterozygosity) and had larger population sizes (as inferred from demographic analyses) relative to eastern populations. Our demographic analyses were, however, equivocal on the role of gene flow in shaping these genetic patterns. Due to the short evolutionary timescale since the *H. destructor* Australian invasion, it is challenging to discern between recent divergence with isolation versus ongoing gene flow. Our draft genome for *H. destructor* is the first assembled for mites in the family Penthaleidae of the order Trombidiformes, an under-represented group with few available reference genomes.

Introduction

The genomes of pest species are providing new evolutionary and applied insights into adaptation to anthropogenic environments. Pesticide resistance in agricultural pests provides an intriguing example of adaptive evolution, where the full genotype-phenotype-environment connection can be made. Toxic pesticides impose selection on populations to evolve phenotypes that can withstand chemical stressors. Because many pesticides exert toxicity through specific biochemical interactions with proteins of target genes, there is a convergence of so called 'target-site' mutations across taxa for mutations that block interactions between pesticides and their target gene (Russell *et al.*, 2004; Dong *et al.*, 2014). Pesticide resistance in agriculture has provided a highly tractable system to study the rise and spread of adaptive variants due to the repeatability of target-site mutations and the ease of characterising adaptive variants (Bass *et al.*, 2015; Feyereisen *et al.*, 2015; Douris *et al.*, 2016). Despite their prevalence in agricultural pests, pesticide resistance does not always occur through the evolution of target-site mutations. Metabolic detoxification (Devonshire & Moores, 1982; Walsh *et al.*, 2018), behavioural responses (Sparks *et al.*, 1989; Zago *et al.*, 2014) and reduced penetrance (Gunning *et al.*, 1991; Puinean *et al.*, 2010) may also contribute to resistance but do not have clear *a priori* genes and (or) alleles. Elucidating these alternative genetic mechanisms of resistance is facilitated with access to adequate reference genomes.

In addition to helping resolve the genomic architecture of adaptation, reference genomes are instrumental in providing more robust demographic inferences. Although reduced representation approaches provide efficient ways of obtaining genomic data for this purpose, their ability to accurately recover the site-frequency spectrum for demographic inference may be dependent on the genomic characteristics of a species, molecular methods implemented and (or) analytical processing (Arnold *et al.*, 2013; Rodríguez-Ezpeleta *et al.*, 2016; Shafer *et al.*, 2016). Reference genomes have aided the investigation of historic demography (Chen *et al.*, 2020; Cao *et al.*, 2022), invasion routes (Rispe *et al.*, 2020; Tay *et al.*, 2022) and patterns of admixture (Zhang *et al.*, 2020) in agricultural pests globally. From a management perspective, understanding demographic processes operating at an individual pest level can be useful for designing

strategies that limit reintroductions from high risk source populations, or for reducing the spread of pesticide resistance (Tay & Gordon, 2019; Thia *et al.*, 2021a).

This study focuses on the redlegged earth mite, *Halotydeus destructor* (Tucker 1925; Penthaleidae, Trombidiformes), a major pest of the grains and pasture industries of Australia. This mite invaded the state of Western Australia from South Africa around 1917 and spread throughout the southern regions of Australia during the 1920s (Newman, 1925; reviewed by Ridsdill-Smith, 1997). *Halotydeus destructor* is a winter-active pest, which diapauses as eggs during the Australian summer (October–March), emerges in autumn (April–May), and typically undergoes its active life stage through to late spring (April–November) (Ridsdill-Smith, 1997). Pyrethroid and organophosphate pesticides have been important in the control of *H. destructor* for many decades (Umina *et al.*, 2019; Arthur *et al.*, 2021). Tolerance to pesticides was first reported in the late 1990s and early 2000s (Hoffmann *et al.*, 1997; Robinson & Hoffmann, 2000), with pyrethroid resistance first detected in field populations in 2006 (Umina, 2007) and organophosphate resistance first detected in 2014 (Umina et al. 2017).

Pyrethroid resistance in *H. destructor* has been relatively straightforward to study due to the substantial contribution of target-site mutations in the voltage-gated sodium channel gene, *para* (Edwards *et al.*, 2018; Cheng *et al.*, 2019; Yang *et al.*, 2020). The frequency of *para* target-site mutations can explain 86% to 91% of the variation in pyrethroid resistance among populations of *H. destructor* (Edwards *et al.*, 2018). Contrastingly, the study of organophosphate resistance has been challenging. Prior research on *H. destructor* has identified mutations in acetylcholinesterase, *ace*, the target gene of organophosphates (Tom Walsh, unpublished data). These mutations include G119S, A201S, and F331Y (*Torpedo califronica* numbering), which have been observed in a variety of other arthropod species exhibiting target-site resistance. However, complicating this picture have been observations of: (1) susceptible mites carrying these mutations; and (2) evidence of *ace* copy number variation in mite populations, based on quantitative PCR (qPCR) analysis. These observations suggest that organophosphate resistance in *H. destructor* may be conferred by a complex combination of target-site mutations, copy number variation in the *ace* gene, and other additional genomic factors. However, the availability of only rudimentary genomic data for *H. destructor* (Edwards *et al.*, 2018) has made follow up investigations challenging.

Historically, populations of *H. destructor* have been considered as showing extensive gene flow across Australia, based on allozyme genetic markers (Weeks *et al.*, 1995; Qin, 1997). Yet, it is also known that isolation-by-distance patterns can manifest over small geographic distances (<2 km) in this mite, based on reduced-representation genomic data (Yang *et al.*, 2020). A long-standing question is whether low levels of genetic differentiation in *H. destructor* are due to high gene flow, large effective population sizes, or a combination of both (Weeks *et al.*, 1995). Determining rates of gene flow among populations of *H. destructor* is pertinent for predictive models of resistance evolution dynamics. Whilst there is evidence that resistance often evolves locally (Yang *et al.*, 2020) and is constrained by climate (Maino *et al.*, 2018), estimates of migration rates across Australia are lacking. Inferring the role of migration in connecting geographically distant populations is becoming increasingly relevant with rising pesticide resistance across Australia (Arthur *et al.*, 2021).

In this work, we present the first draft genome for *H. destructor*. Our genome assembly reveals a substantial amplification of *ace* in *H. destructor*, which may have evolved in response to selection by organophosphate pesticides. Comparative analysis revealed this amplification to be unusual with respect to a range of other arthropods, although another mite from the Trombidiformes, the two-spotted mite, *Tetranychus urticae*, has also exhibited an extensive amplification of *ace* in its genome. Population genomic analyses indicated that *H. destructor* populations in Australia are regionally structured and contain unique compositions of *ace* alleles.

Demographic inference from genome-wide joint site-frequency spectra (SFS) were equivocal on the role of gene flow between eastern and western populations, although a scenario of isolation was the most parsimonious. Our draft genome for *H. destructor* is an important resource, aiding in ongoing efforts to characterise the genomic architecture of pesticide resistance and other adaptive processes in this pest.

Methods

De novo genome assembly and annotation

Biological material and sequencing — To assemble the nuclear genome, samples of adult *H. destructor* were collected from a grassland in Wantirna South, Victoria, Australia (−37°52'11.6", 145°11'46.1"), in June 2020. Collections were via suction using a blower vacuum with a fine gauze mesh placed over the end of the vacuum tube. Mites were placed into a plastic container with paper towel and vegetation for transportation. Once in the laboratory, *H. destructor* individuals were morphologically identified by light microscopy and separated from other invertebrates. DNA was extracted from pooled individuals (n = 70) using a DNeasy® Blood & Tissue kit (Qiagen, Hilden, Germany). A continuous long-read library was prepared and sequenced using PacBio Sequel II technology by Berry Genomics (Berry Genomics Co. Ltd, Beijing, China). PacBio sequencing yielded 761,202 reads, with a mean of 13,106 (±6,705 SD) bp, a maximum of 20,938 bp, with ~78% of reads having a length of ≥10,000 bp. Additionally, a 150 bp paired-end Illumina short-read library was prepared and sequenced (Novogene Co. Ltd, Hong Kong), which yielded 20,748,875 reads.

We employed RNA sequence data from a prior study of *H. destructor* (Tom Walsh, unpublished data) for evidence-based prediction of genes in our draft genome assembly. These transcriptomes came from two populations in Western Australia, one from Capel and one from Manjimup, both of which were collected in 2016 and shown to be resistant to organophosphates through standard pesticide bioassays (Umina *et al.*, 2017). RNA was extracted using Trizol (Life Sciences, Carlsbad, CA, USA), transcriptome libraries were prepared using a custom protocol, and high throughput Illumina 75 bp paired-end sequencing was performed at the Australian National University (Canberra, Australian Capital Territory, Australia). These samples were quality trimmed with TRIMMOMATIC. In total, 20,183,114 and 22,688,705 paired-end reads were available for the Capel and Manjimup samples, respectively.

Genome assembly — Raw Illumina reads were analysed with JELLYFISH (Marcais & Kingsford, 2012) and GENOMESCOPE (Vurture *et al.*, 2017) to perform preliminary genomic profiling in *H. destructor*. Using a kmer size of 21, we inferred a genome size of ~50 Mb, with a good fit between the observed and expected kmer distribution (98%). The heterozygosity rate (the proportion of polymorphic sites relative to total sites) was estimated at 1.4%. Raw PacBio reads were assembled with SMARTDENOVO (Liu *et al.*, 2021) to yield a preliminary set of contigs, with a minimum read length threshold of 10,000 bp. We then checked for contamination and polished the genome using an established pipeline (Korhonen *et al.*, 2019). For polishing the assembly, Illumina short-reads were quality trimmed using FASTP (Chen *et al.*, 2018).

Genome annotation — Gene prediction and functional annotation was accomplished using an AUGUSTUS software-based pipeline (Korhonen *et al.*, 2022), which employs RNA-seq data as evidence data to assist gene prediction. For the annotation of predicted amino acid sequences, INTERPROSCAN was used to search the InterPro domain database (Zdobnov & Apweiler, 2001) and BLASTP (threshold *E*-value: ≤10−8) was employed to sequence homology databases Swiss-Prot (Magrane, 2011; accessed 30 June 2021), KEGG (Kanehisa *et al.*, 2012; accessed 30 June 2021) and NCBI NR (Pruitt *et al.*, 2012; accessed 4 February 2021). BUSCO analysis

(Simão *et al.*, 2015) was used to assess the completeness of our protein annotations. Proteins were screened against the core arthropod ("arthropoda_odb10") and the arachnid ("arachnida_odb10") gene sets.

Not all proteins (6,518 sequences) were confidently assigned a GO term. We manually annotated the *ace* copy variants and performed a supplementary GO term assignment with BLAST2GO (Conesa *et al.*, 2005; Conesa & Götz, 2008; Götz *et al.*, 2008). All proteins reported for the subclass, Acari, in the NCBI database were downloaded to construct a local BLAST database and mapped against our unannotated genes with BLAST2GO using the default parameters. Collectively, 2,089 additional genes were assigned GO terms. These supplementary annotations are not reported in our NCBI submission but can be found in supplementary Appendix 1.

Comparative analysis of *ace* gene sequences

Following annotation, it was apparent that the genome of *H. destructor* harbours many copies of the *ace* gene. We performed a comparative analysis to understand how copy number variation in *ace* varies among other arthropod genomes. *Bemisia tabaci* (Silverleaf whitefly, GenBank GCF_001854935.1), *Leptinotarsa decemlineata* (Colorado potato beetle, GenBank GCF_000500325.1), *Myzus persicae* (green peach aphid, GenBank GCF_001856785.1), *Plutella xylostella* (diamondback moth, GenBank GCF_905116875.1) and *T. urticae* (GenBank GCA_000239435.1) were selected as major agricultural arthropod pests that have evolved resistance to multiple pesticides. *Leptotrombidium deliense* (a species of chigger mite, GenBank GCA_003675905.2) was used as a representative tetranychid mite that is a not an agricultural pest. *Drosophila melanogaster* (GenBank GCA_000001215.4) was also included. Our arthropod gene trees were rooted on the *Torpedo californica ace* sequence (Pacific electric ray, GenBank X03439), which is the canonical reference for the numbering of *ace* target-site mutations.

Downstream analyses were performed in R. First, we constructed a multispecies gene tree to contrast the *H. destructor ace* genes to those in other species. We imported *ace* protein sequences with the packages BIOSTRINGS (Pagès *et al.*, 2017) and APE (Paradis *et al.*, 2004). Due to truncation and variable length of protein products among *ace* genes, we restricted the multispecies analysis to *ace* genes with proteins ≥500 amino acids in length. ClustalW alignment of *ace* protein sequences was performed with the MSA package (Bodenhofer *et al.*, 2015). Following alignment, amino acid sites with missing data were removed, leaving 207 sites for gene tree construction. A distance matrix among protein sequences was derived using the number of amino acid differences. Because we analysed a mix of orthologous and paralogous *ace* sequences in this multispecies analysis, we chose not to implement a model of protein evolution when deriving the distance matrix. We used the APE package to construct a neighbour joining tree from the distance matrix, with node support estimated with 500 bootstrap replicates.

We constructed a second gene tree to focus on relationships among *ace* genes within *H. destructor*. In this case, we chose to examine all *ace* genes, including those with short, truncated protein sequences <500 amino acids in length. Hence, the distance matrix was derived using pairwise deletion among *ace* protein sequences. Again, a neighbour joining tree was constructed using the number of amino acid differences, with 500 bootstrap replicates to estimate node support. The R package GENOMALICIOUS (Thia & Riginos, 2019) was used to visualise relative gene positions and structural variation among select genes.

Population genomics

Biological material and sequencing — We used our assembled genome to characterise population genetic patterns among Australian *H. destructor* populations. Samples for population genomic analyses came from a study investigating the mechanisms of organophosphate resistance (Table S1). These samples came from

populations in the states of Western Australia (Manjimup, Tambellup, Treeton, and Yongarillup) and Victoria (Colbinabbin Res, Colbinabbin Sus; "Res" for resistant, and "Sus" for susceptible to organophosphates). Briefly, DNA from these mites was extracted using a DNeasy® Blood and Tissue Kit (Qiagen, Hilden, Germany) and sequenced using paired-end Illumina 150 bp (Novogene). These data were combined with the pool-seq data from the Wantirna South population used for the genome assembly.

Bioinformatics — Reads from each pool-seq library were quality trimmed using FASTP and truncated to 100 bp; trimmed reads <100 bp were discarded. After trimming, the mean number of reads per library was 29,570,401 (±3,351,748 SD). Trimmed read pairs were mapped to our draft genome with BOWTIE2 using a global alignment and the "--sensitive" mapping option. SAMTOOLS (Li *et al.*, 2009; Danecek *et al.*, 2021) was used to discard reads with mapping quality <30 and those that were found to be PCR duplicates. The BAM files from pool-seq libraries from the same population were merged with SAMTOOLS, and we refer to these merged libraries herein as our "population samples". In preliminary analyses, we observed considerable variance in depth and coverage of contigs that were <0.2 Mb (Figure S1). Therefore, we limited our reference to those contigs ≥0.2 Mb for population genomic analyses, with a focal set of 56 long contigs.

Variant calling was performed on population sample BAM files using FREEBAYES (Garrison & Marth, 2012) with the arguments: "--min-alternate-count 3" and "--pooled-continuous". The FREEBAYES script, FASTA_GENERATE_REGIONS.PY, was used to segment the genome for variant calling in parallel. We used BCFTOOLS (Li, 2011; Danecek *et al.*, 2021) to combine VCF files from across the parallel variant calling runs. Variant filtering was performed with VCFTOOLS (Danecek *et al.*, 2011) to retain only single nucleotide polymorphisms (SNPs; no indels) with a minor allele count of 3, a minor allele frequency of 0.01, minimum quality of 30, minimum depth of 20 reads per sample, no missing data, and to exclude SNPs from repetitive regions of the genome. There were 599,969 SNPs remaining after filtering. The VCF file was imported into R using the package POOLFSTAT (Hivert *et al.*, 2018).

We calculated the population genomic heterozygosity, π, using the estimator $\hat{\pi}$ outlined in Ferretti, Ramos-Onsins, & Pérez-Enciso (2013):

$$\hat{\pi} = \frac{n}{n-1}\frac{1}{L}\sum_{i}\frac{2}{r_i(r_i-1)}\sum_{m=1}^{r_i-1} m(r_i-m)\xi_{im}$$

Here, n is the number of haploid individuals in a pool, L is the length of a focal genomic window, i is the index of a focal "cluster" nested within a genomic window, r_i is the read depth within cluster i, m is the observed alternate read count (ranging from 1 to $r_i - 1$), and ξ_{im} is the number of sites within cluster i with m counts of the alternate allele. The original formulation of this estimator considers a "cluster" to be a region within a genome with the same read depth, which could be ≥1 bp. For clusters >1 bp, the iteration from $m = 1$ to $r_i - 1$ for the summation of $m(r_i - 1)\xi_{im}$ provides a reflection of the site frequency spectrum across multiple adjacent sites with the same read depth. We simplified the calculation by defining clusters within genomic windows as single discrete nucleotide positions (1 bp) that were either polymorphic (a SNP) or monomorphic. This simplifies the calculation because there is a single value of m for the ith cluster, m_i, and the value of ξ_{im} is always 1. And therefore:

$$\hat{\pi} = \frac{n}{n-1}\frac{1}{L}\sum_{i}\frac{2(m_i(r_i-m_i))}{r_i(r_i-1)}$$

We calculated $\hat{\pi}$ for non-overlapping 5,000 bp windows for our focal set of 56 contigs >0.2 Mb in length. We calculated $\hat{\pi}$ within genomic windows because sequencing depth was not adequate across the entire length

of all contigs. We obtained the mean contig $\hat{\pi}$ by averaging across windows within each contig, and the mean genome-wide $\hat{\pi}$ was calculated as the mean of mean contig $\hat{\pi}$ (a "mean of means"). To test the hypothesis that populations varied in their levels of heterozygosity, we constructed a mixed effects model fitting the mean contig $\hat{\pi}$ as a function of the fixed "population" effect and the random effect of "contig". The Wantirna South population was used as the reference level (intercept). Our model was tested using the LME4 (Bates *et al.*, 2014) and LMERTEST (Kuznetsova *et al.*, 2017) R packages. The MULTCOMPVIEW R package (Graves *et al.*, 2019) was used to delineate groups with non-significantly different $\hat{\pi}$ in pairwise *post hoc* comparisons.

Pairwise genome-wide F_{ST} was calculated between all population sample combinations using all available SNPs, with the POOLFSTAT package. The POOLFSTAT calculation of F_{ST} uses the expected allele frequencies from observed read counts to account for two sources of sampling variance inherent to pool-seq data: estimation of the population allele frequency and estimation of the pooled sample allele frequency. This F_{ST} corrected for pool-seq sampling variance is analogous to Weir & Cockerham's (1984) ϑ and is calculated as:

$$F_{ST} = \frac{Q_1 - Q_2}{1 - Q_2}$$

Here, Q_1 is the identity in state probability for two genes sampled within a population pool, and Q_2 is the identity in state probability for two genes sampled between two population pools.

In addition, we used a principal component analysis (PCA) to visualise patterns of population structure. The PRCOMP function in R was used to conduct a PCA on allele frequency covariances, using the arguments "scale=FALSE, center=TRUE". We performed two PCAs: one on genome-wide SNPs, and another exclusively on SNPs residing in *ace* genes.

Demographic inference

To test the importance of gene flow on regional population structure in *H. destructor*, we used forward in time demographic simulations, implemented in the program *dadi* (Gutenkunst *et al.*, 2009). To model the invasion of *H. destructor* into Australia, we included an ancestral change in population size to represent the bottleneck in the original founding population, a split of the founding population into an eastern and western population, and changes in population size of each population following the east–west split. We then modelled three different scenarios of gene flow: (1) asymmetric migration, where the magnitude of gene flow could vary east-to-west or west-to-east; (2) symmetric migration, where gene flow between regions was equivalent; and (3) isolation, where no gene flow occurred between regions.

Demographic parameters modelled in *dadi* are relative parameters, scaled by N_{anc}, the ancestral population size. In our models, the parameter T_1 represented the timing of the ancestral bottleneck in the founding Australian population, with size v_{anc}. The parameter T_2 represented the scaled time of the split of the ancestral population into a western and eastern population, each with size v_{west} and v_{east}. The time parameters were scaled relative to the ancestral population size at the xth generation, where $T_x = t_x/(2N_{anc})$, with t_x equal to the number of actual generations that have passed. The population size parameters represent the proportional size of the ith population relative to the original ancestral population, such that $N_i = v_i N_{anc}$, where N_i is the absolute size of the ith population. Finally, gene flow between regional populations was modelled with the scaled migration parameter $M_{i,j}$, the number of haploid genomes that move from the jth population into the ith population per generation. For the isolation scenario, we did not model a migration parameter. For the symmetric migration scenario, gene flow was equivalent in both directions, so we modelled the single parameter, M. For the asymmetric migration scenario, we allowed the direction of gene flow to vary by fitting two migration parameters, $M_{east,west}$ (west-to-east gene flow) and $M_{west,east}$ (east-to-

west gene flow). The scaled migration parameters are related to the proportion of migrants per generation, $m_{i,j}$, through $M_{i,j} = m_{i,j}N_{anc}$.

To infer the values for demographic parameters, we established three eastern versus western population pairs with the following expected haploid genomes: Wantirna South and Manjimup, n =140 and 112; Treeton and Colbinabbin Res, n = 84 and 100; and Yongarillup and Colbinabbin Sus, n = 164 and 100. SNPs were thinned to one randomly sampled SNP per 5,000 bp. Pooled allele frequencies of thinned SNPs were converted into expected joint site-frequency spectra (SFS) using the DADI_INPUTS_POOLS function from the GENOMALICIOUS R package, using the argument "methodSFS='probs'" to infer allele counts under binomial expectations. Prior to demographic analyses, the joint SFS were downward projected to 20-by-20 haploid genomes with *dadi*, with final analysed SNP numbers as follows: Wantirna South and Manjimup, 8,275 SNPs; Treeton and Colbinabbin Res, 8,218 SNPs; and Yongarillup and Colbinabbin Sus, 8,248 SNPs.

Demographic models were fit using *dadi*'s "log_fmin" optimiser, in two phases. The first phase was used to explore parameter space, and the second phase homed in on the most likely parameters. We used ΔAIC criteria to evaluate model support, whereby a pair of models were considered comparable if ΔAIC > –10, and ΔAIC = best model AIC – alternate model AIC (Burnham & Anderson, 2004). We note that this is a conservative threshold for assessing model comparability. We assessed model comparability with ΔAIC at two levels: (1) *within* scenarios, *within* population pairs, to assess the range of parameter values across replicate simulations with similar likelihoods; and (2) *between* scenarios, *within* population pairs, to test support for different demographic scenarios.

In the first phase, we executed 200 replicate simulations, each with 500 maximum iterations, for each scenario and population pair. The seeding parameter values were all set to 0.1. Lower limits for all parameters were set to 1e–10. Upper limits for T_1, T_2, ν_{anc}, ν_{east}, and ν_{west} were set to 1. Upper limits for M, $M_{east,west}$, and $M_{west,east}$ were set to 10. For each replicate simulation, seeding parameters were perturbed by a factor of 5 within the limits. At the end of the first phase, we identified a set of comparable replicate simulations within scenarios, within population pairs, that had similarly high likelihoods (ΔAIC > –10 relative to the best replicate simulation). The median parameter values across this set of comparable replicate simulations were used to seed parameters in the second phase. In the second phase, we executed 100 replicate simulations, each with 500 maximum iterations. Seeding parameters were perturbed by a factor of 2 and limits were the same as those in the first phase. At the end of the second phase, we again identified the best replicate simulation and those comparable replicate simulations with similarly high likelihoods.

The best replicate simulation from the second phase was used in comparisons between scenarios, within population pairs, as per standard practise. Comparable replicate simulations from the second phase were used to assess the distributions and correlations of parameter values across optimisation runs that achieved similar likelihoods, within scenarios, within population pairs. We note that the primary motivation of our demographic analyses was to try and tease out the importance of migration between eastern and western populations. We did not use these models to infer biologically meaningful parameters, which would require more extensive simulations and genomic bootstrapping (Gutenkunst *et al.*, 2009).

Results

De novo genome assembly and annotation

Our draft *de novo* assembly of the *H. destructor* genome comprised 132 contigs, with a total length of 48.90 Mb, an N50 of 874,492, an L50 of 14, and GC content of 45.3%. Our annotation of this genome predicted 14,583 genes. BUSCO analyses indicated that our genome assembly captured a substantial proportion of the core genes expected for arthropods and arachnids. Of the 1,013 single copy orthologs in arthropods, 906 (89.5%) were identified, 892 (88.1%) as single copies and 14 (1.4%) as duplicates; 30 (3.0%) were fragmented and 77 (7.5%) were missing. Of the 2,934 single copy orthologs in arachnids, 2,641 (90%) were found, 2,571 (87.6%) as single copies and 70 (2.4%) as duplicates; 34 (1.2%) were fragmented and 259 (8.8%) were missing. To further assess the quality of the genome, we compared our assembly against all available genomes for trombidiform mites, with respect to the number of contigs relative to genome size (Figure 1). This comparison demonstrates that our assembly is consistent with other genomes published for mites of the order Trombidiformes.

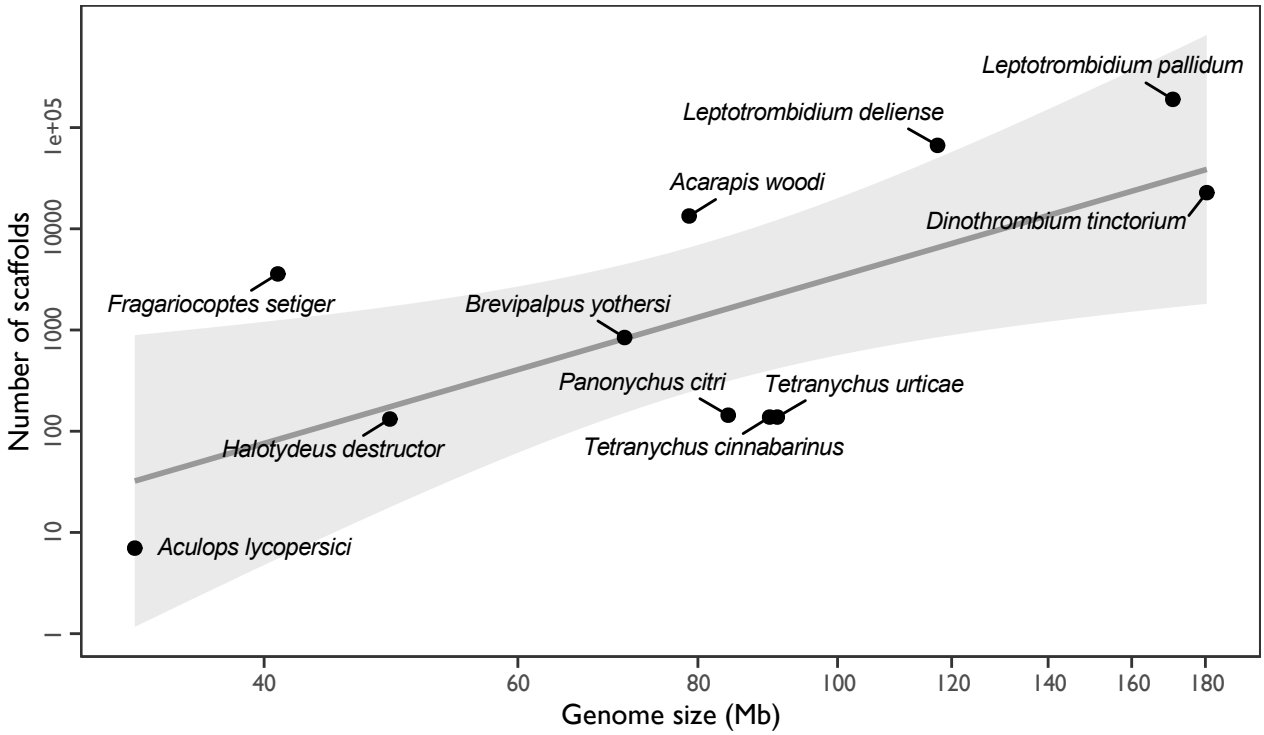

Figure 1. Genome assembly comparison among mites from the order, Trombidiformes. The *x*-axis is the assembled genome size, and the *y*-axis is the number of scaffolds, displayed here on a $\log_{10}$ scale. The dark grey line is the linear fit, with light grey shading representing the standard error of predicted values. Each point represents a genome from NCBI: *Acarapis woodi*, GCA_023170135.1; *Aculops lycopersici*, GCA_015350385.1; *Brevipalpus yothersi*, GCA_003956705.1; *Dinothrombium tinctorium*, GCA_003675995.1; *Fragariocoptes setiger*, GCA_019968535.1; *Halotydeus destructor*, GCA_022750525.1 (from this study); *Leptotrombidium deliense*; GCA_003675905.2*; Leptotrombidium pallidum*, GCA_013372475.1; *Panonychus citri*, GCA_014898815.1; *Tetranychus cinnabarinus*, GCA_022266195.1; and *Tetranychus urticae*, GCF_000239435.1.

Comparative analysis of *ace* gene sequences

In total, 28 *ace* genes were identified in the present draft genome for *H. destructor*. The number of *ace* genes in *H. destructor* were markedly higher than that observed in the other arthropod species surveyed, only exceeded by *T. urticae* (46 genes), whereas others had far lower numbers: *B. tabaci* (4 genes), *D. melanogaster* (1), *L. decemlineata* (4), *L. deliense* (10), *M. persicae* (3), and *P. xylostella* (4) (Figure 2). Collectively, the three mites from the order Trombidiformes (*H. destructor*, *L. deliense* and *T. urticae*) had

more *ace* genes relative to arthropods in different orders. Most of the *ace* genes in *H. destructor* and *T. urticae* were ≥500 residues in length (25 and 40 *ace* genes), whereas only 2 genes in *L. deliense* were ≥500 residues long (Figure 2).

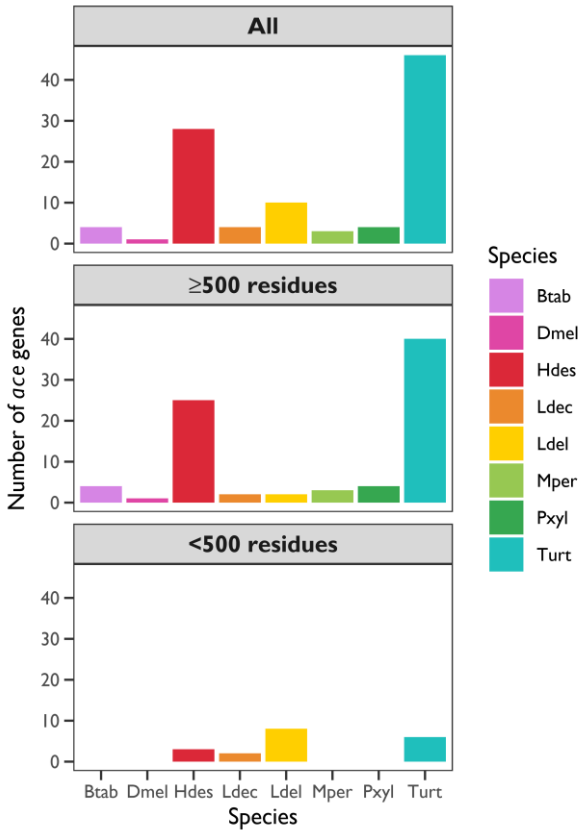

Figure 2. Counts of annotated *ace* genes in select arthropod genomes. Species are on the *x*-axis and the number of annotated *ace* genes in each genome are on the *y*-axis. Bars are coloured by species (see legend). Each panel contains the counts of *ace* genes for different length thresholds: top, all *ace* genes; middle, *ace* genes with a length ≥500 amino acid residues; bottom, *ace* genes with a length <500 amino acid residues. Species abbreviations: "Btab", *Bemisia tabaci*; "Dmel", *Drosophila melanogaster*; "Hdes", *Halotydeus destructor*; "Ldec", *Leptinotarsa decemlineata*; "Ldel", *Leptotrombidium deliense*; "Mper", *Myzus persicae*; "Pxyl", *Plutella xylostella*; "Turt", *Tetranycus urticae.*

Based on our multi-species neighbour joining tree (Figures 3 & S2), the most ancestral copy of *ace* in the *H. destructor* genome was HDE_14279. Notably, there was a large cluster of 17 *H. destructor ace* genes with shared homology, suggesting a large expansion event from a single gene copy. Such an expansion was also evident in *T. urticae*, with two distinct clusters of *ace* genes present. Homologues of the remaining copies of the *H. destructor ace* genes were present in other species (Figure 3). There was evidence that some of the *H. destructor ace* gene copies may be associated with structural rearrangements. The clearest example being the *ace* genes found on contig 344 and contig 1156, which exhibit very similar proteins sequences relative to other copies (Figures 3 & S2) and appear to be associated with an inversion and deletion event (Figure S4).

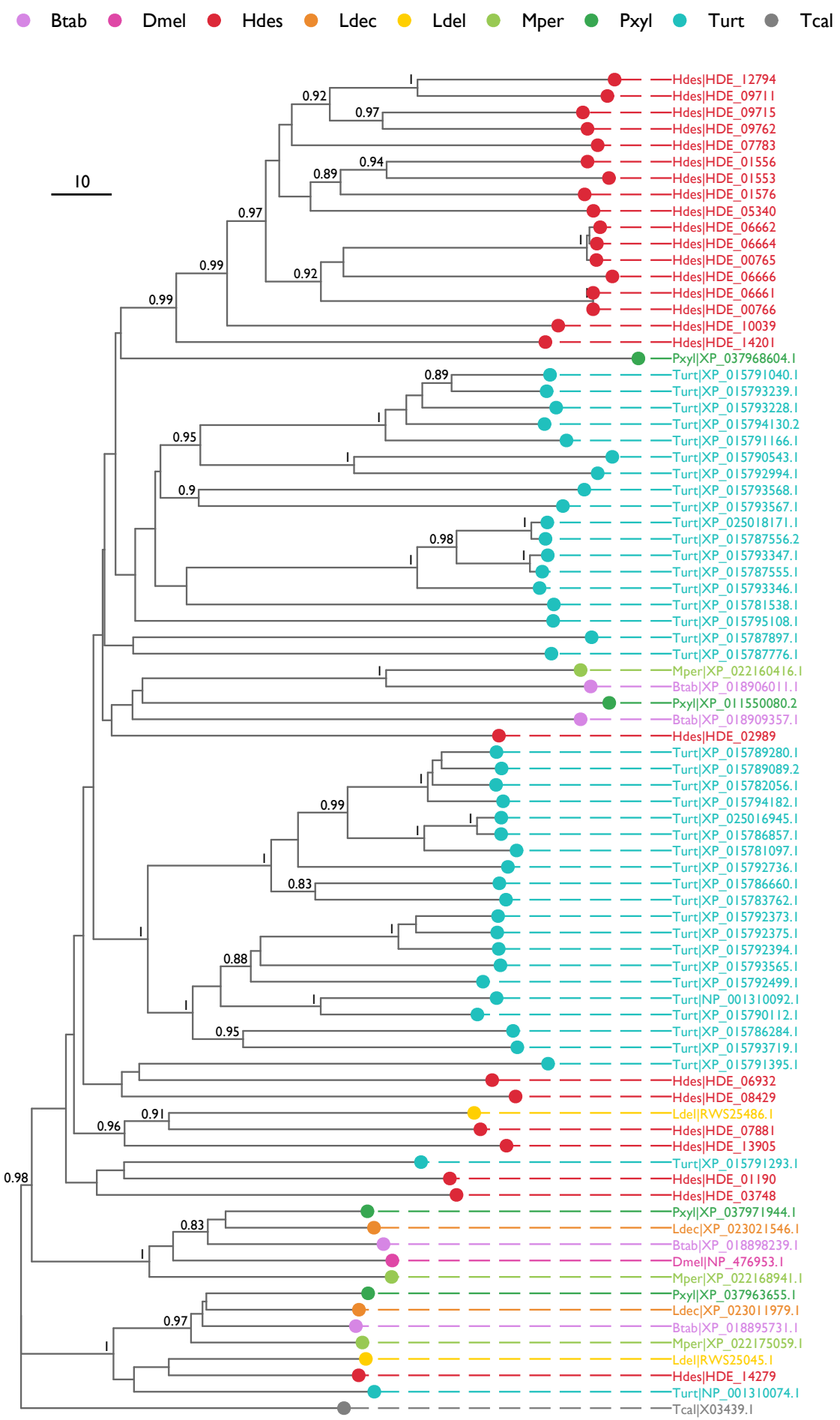
Btab
Dmel
Hdes
Ldec
Ldel
Mper
Pxyl
Turt
Tcal
10
Hdes|HDE_12794
Hdes|HDE_09711
Hdes|HDE_09715
Hdes|HDE_09762
Hdes|HDE_07783
Hdes|HDE_01556
Hdes|HDE_01553
Hdes|HDE_01576
Hdes|HDE_05340
Hdes|HDE_06662
Hdes|HDE_06664
Hdes|HDE_00765
Hdes|HDE_06666
Hdes|HDE_06661
Hdes|HDE_00766
Hdes|HDE_10039
Hdes|HDE_14201
Pxyl|XP_037968604.1
Turt|XP_015791040.1
Turt|XP_015793239.1
Turt|XP_015793228.1
Turt|XP_015794130.2
Turt|XP_015791166.1
Turt|XP_015790543.1
Turt|XP_015792994.1
Turt|XP_015793568.1
Turt|XP_015793567.1
Turt|XP_025018171.1
Turt|XP_015787556.2
Turt|XP_015793347.1
Turt|XP_015787555.1
Turt|XP_015793346.1
Turt|XP_015781538.1
Turt|XP_015795108.1
Turt|XP_015787897.1
Turt|XP_015787776.1
Mper|XP_022160416.1
Btab|XP_018906011.1
Pxyl|XP_011550080.2
Btab|XP_018909357.1
Hdes|HDE_02989
Turt|XP_015789280.1
Turt|XP_015789089.2
Turt|XP_015782056.1
Turt|XP_015794182.1
Turt|XP_025016945.1
Turt|XP_015786857.1
Turt|XP_015781097.1
Turt|XP_015792736.1
Turt|XP_015786660.1
Turt|XP_015783762.1
Turt|XP_015792373.1
Turt|XP_015792375.1
Turt|XP_015792394.1
Turt|XP_015793565.1
Turt|XP_015792499.1
Turt|NP_001310092.1
Turt|XP_015790112.1
Turt|XP_015786284.1
Turt|XP_015793719.1
Turt|XP_015791395.1
Hdes|HDE_06932
Hdes|HDE_08429
Ldel|RWS25486.1
Hdes|HDE_07881
Hdes|HDE_13905
Turt|XP_015791293.1
Hdes|HDE_01190
Hdes|HDE_03748
Pxyl|XP_037971944.1
Ldec|XP_023021546.1
Btab|XP_018898239.1
Dmel|NP_476953.1
Mper|XP_022168941.1
Pxyl|XP_037963655.1
Ldec|XP_023011979.1
Btab|XP_018895731.1
Mper|XP_022175059.1
Ldel|RWS25045.1
Hdes|HDE_14279
Turt|NP_001310074.1
Tcal|X03439.1

Figure 3. Neighbour joining tree of *ace* protein sequences from select arthropod species. Each tip represents an annotated *ace* gene, with points and labels coloured by species (see legend). The *ace* sequence from *Torpedo californica* ("Tcal") was used to root the tree. Species abbreviations: "Btab", *Bemisia tabaci*; "Dmel", *Drosophila melanogaster*; "Hdes", *Halotydeus destructor*; "Ldec", *Leptinotarsa decemlineata*; "Ldel", *Leptotrombidium deliense*; "Mper", *Myzus persicae*; "Pxyl", *Plutella xylostella*; "Tcal", *Torpedo californica*; "Turt", *Tetranycus urticae*. The scale bar indicates the number of amino acid differences. Nodes labels indicate ≥80% bootstrap support.

Population genomics

We investigated patterns of genome-wide and *ace*-specific population structure within *H. destructor* from the states of Western Australia (western populations) and Victoria (eastern populations) (Figure 4a). The mean genomic heterozygosity, $\hat{\pi}$, ranged from 0.00416 to 0.00446 (Figure 4b). There were statistically significant differences among populations (χ^2 = 2,651, df = 7, $p < 0.001$). Overall, eastern populations had lower genome-wide genomic heterozygosity than western populations. Within Victoria, Wantirna South had the greatest diversity ($\hat{\pi}$ = 0.00431), followed by Colbinabbin Res ($\hat{\pi}$ = 0.00422), with the least diversity in Colbinabbin Sus ($\hat{\pi}$ = 0.00416). Within Western Australia, genomic heterozygosity was largely comparable among Yongarillup ($\hat{\pi}$ = 0.00446), Manjimup ($\hat{\pi}$ = 0.00445), Treeton ($\hat{\pi}$ = 0.00442), and Tambellup ($\hat{\pi}$ = 0.00422). Estimates of genomic heterozygosity were relatively consistent across populations for a given genomic contig (Figure S3).

Australian populations of *H. destructor* exhibited a pattern of hierarchical regional population structure (Figure 4c). Pairwise genome-wide F_{ST} values ranged from 0.006 to 0.069; they were highest between states ($0.032 \leq F_{ST} \leq 0.068$), while they were lower within Western Australia ($0.014 \leq F_{ST} \leq 0.039$) and within Victoria ($0.006 \leq F_{ST} \leq 0.015$). These data suggest that populations from Western Australia are genetically distinct from those in Victoria, whereas there was little variation among populations within Victoria and some structuring within Western Australia. Within Western Australia, Tambellup was the most differentiated from the other populations. These patterns of differentiation were recapitulated in the PCA of genome-wide SNP allele frequencies (Figure 4d), with regions separated on PC1 and variation within Western Australia separated on PC2. PCA of allele frequencies for *ace* gene SNPs reflected genome-wide patterns, indicating that the local populations contain unique compositions of *ace* variants (Figure 4e).

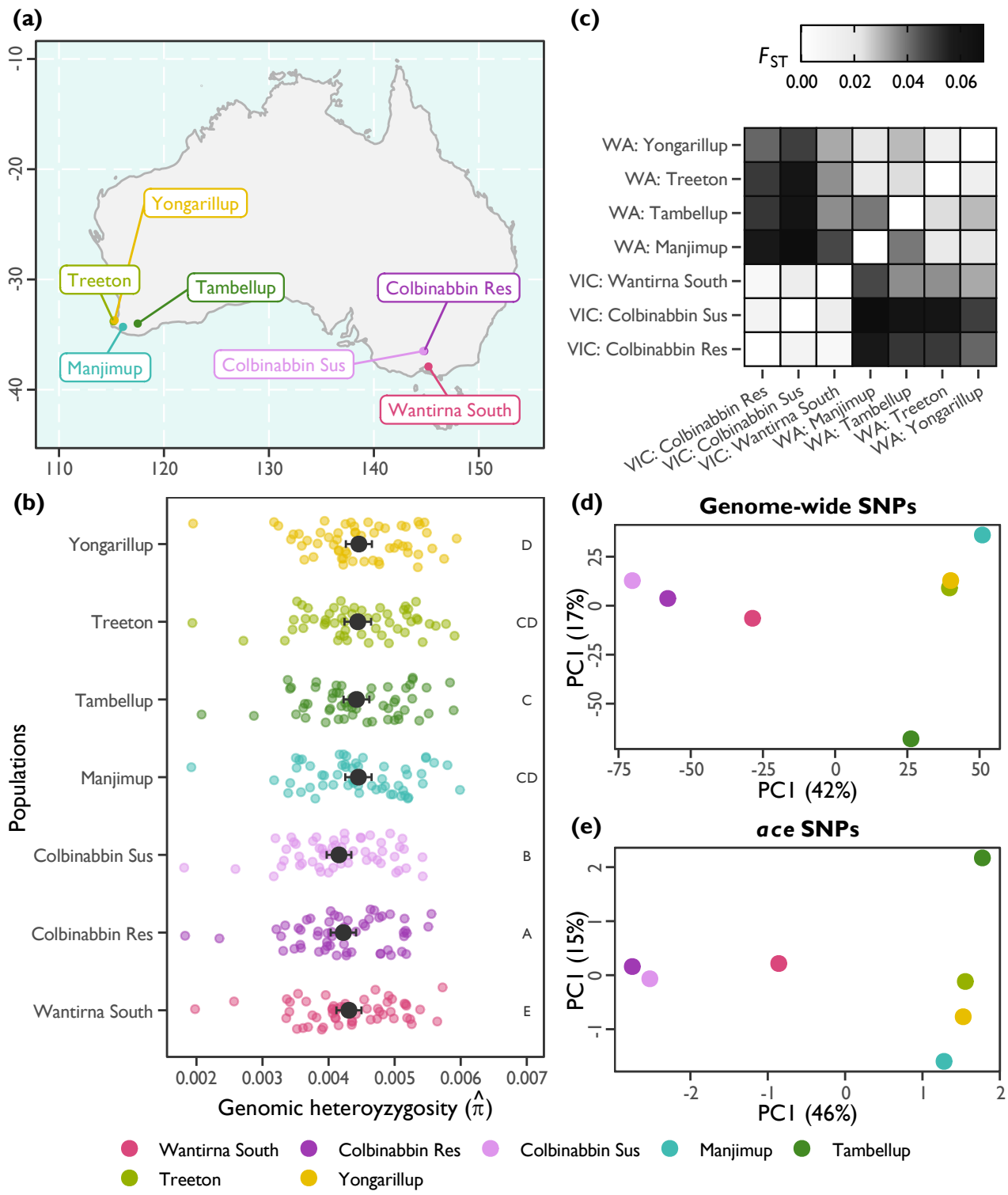

Figure 4. Population genetics of Australian *H. destructor*. (a) Sampling locations of mite populations in Australia: the Manjimup, Treeton, Tembellup, and Yongarillup populations were from state of Western Australia, and the Colbinabbin (Res and Sus) and Wantirna South populations were from the state of Victoria. (b) Population genomic heterozygosity. Populations are on the *x*-axis, with genomic heteroygosity (π) on the *y*-axis. Coloured points represent the average π for each contig in our draft genome assembly. Large black points represent the average $\hat{\pi}$ across contigs with 95% confidence intervals. Capital letters indicate groups with non-significant differences in genomic heterozygosity. (c) Pairwise genome-wide F_{ST} values among all populations. Darker shading indicates a greater genetic differentiation (see legend). State abbreviations: 'VIC', Victoria; 'WA', Western Australia. (d) Principal component analysis (PCA) of genome-wide allele frequencies. (e) PCA of *ace* gene allele frequencies. (d,e) Points represent a sample, with colour representing the population (see legend). The *x*-axes and *y*-axes represent the first and second principal component, respectively, with numbers in parenthases indicating the percent of explained variance in allele frequencies.

Demographic inference

We used the forward in time simulator, *dadi*, to infer demographic parameters under three modelled scenarios: asymmetric migration, symmetric migration, and isolation (Figure 5a). By our second phase of optimisation, most of the 500 replicate simulations within scenarios and population pairs had reached comparable regions of likelihood space, with ΔAIC > –10 relative to the best model. For the asymmetric migration, symmetric migration, and isolation scenarios (respectively), the number of comparable replicate simulations in phase 2 were as follows: Manjimup and Wantirna South, n = 99, 99, and 98; Treeton and Colbinabbin Res, n = 99, 99, 98; Yongarillup and Colbinabbin Sus, n = 99, 99, and 98 (out of a total of 100).

The best replicate simulations for each scenario and population pair are reported in Table 1. Within population pairs, log-likelihoods of the best replicate simulation were virtually identical for all three scenarios. Across population pairs, and across scenarios, values of T_1, v_{anc}, T_2, v_{west} and v_{east} were very small, likely reflecting the recent invasion history of *H. destructor* into Australia and the small sizes of the invasive populations. Western populations had slightly larger inferred population sizes (2.0e–9 ≤ v_{west} ≤ 5.6e–7) relative to eastern populations (7.3e–10 ≤ v_{east} ≤ 2.7e–7).

Although the best supported (lowest AIC) scenario was always the isolation scenario, the symmetric and asymmetric migration scenarios produced comparable fits to the observed SFS (Table 1). In all population pairs, the symmetric migration scenario had ΔAIC = –2, and the asymmetric migration scenario had a ΔAIC = –4, reflecting the cost of adding one and two additional migration parameters, respectively. Indeed, the modelled SFS were indistinguishable among scenarios (Figure 5b) and had identical residual patterns (Figures S5–S7). For the best replicate simulations, estimates of M, $M_{west,east}$, and $M_{east,west}$ were ≤1 (Table 1). However, examination of the parameter estimates across comparable replicate simulations (ΔAIC > −10, relative to best replicate simulation) within each scenario revealed diverse optima under the symmetric and asymmetric migration scenarios. Whilst timing and size parameters exhibited tight convergence across comparable replicate simulations (Figures S8–S10), migration parameters were more diffuse (Figures S8 & S9). Migration rate estimates were not associated with the timing of the ancestral bottleneck, the east–west split, or population sizes (Figures S11 & S12). Furthermore, asymmetric migration rate parameters were uncorrelated; increasing gene flow in one direction was not related to decreasing gene flow in the alternate direction (Figures S11 & S12). Our results suggest that inferring the amount and directionality of gene flow in *H. destructor* is challenging because many values of M, or $M_{east,west}$ and $M_{west,east}$ produce comparable fits to the observed SFS within their respective scenarios.

Table 1. Log-likelihood, AIC, ΔAIC, and demographic parameters for the best replicate simulations for each population pair and scenario. Population abbreviations: 'Manj', Manjimup; 'WanS', Wantirna South; 'Tree', Treeton; 'ColR', Colbinabbin Res; 'Yong', Yongarillup; 'ColS', Colbinabbin Sus.

Population	Scenario	Log-likelihood	AIC	ΔAIC	T_1	v_{anc}	T_2	v_{west}	v_{east}	M	$M_{west,east}$	$M_{east,west}$
Manj/WanS	Asymmetric	–608.1	1230.3	–4	1.0e–10	4.8e–10	1.0e–10	1.8e–9	1.1e–9		0.75	0.22
	Symmetric	–608.1	1228.3	–2	1.0e–10	4.7e–10	1.0e–10	1.8e–9	1.1e–9	2.7e–1		
	Isolation	–608.1	1226.3	0	1.5e–9	7.1e–9	1.2e–10	2.1e–9	1.3e–9			
Tree/ColR	Asymmetric	–611.7	1237.5	–4	3.0e–10	1.4e–9	5.3e–10	9.2e–9	5.0e–9		0.26	1.7e–2
	Symmetric	–611.7	1235.5	–2	1.1e–10	5.3e–10	1.3e–9	2.3e–8	1.2e–8	1.0		
	Isolation	–611.7	1233.5	0	3.7e–10	1.8e–9	1.2e–10	2.1e–9	1.1e–9			
Yong/ColS	Asymmetric	–632.1	1278.3	–4	1.4e–10	6.6e–10	1.0e–10	2.0e–9	7.3e–10		4.0e–3	0.12
	Symmetric	–632.1	1276.3	–2	4.7e–8	2.3e–7	2.9e–8	5.7e–7	2.1e–7	7.5e–1		
	Isolation	–632.1	1274.3	0	2.0e–10	9.7e–10	1.1e–10	2.1e–9	7.9e–10			

Note: See supplementary Figures S8–S13 for distributions and correlations of parameter estimates across comparable replicate simulations.

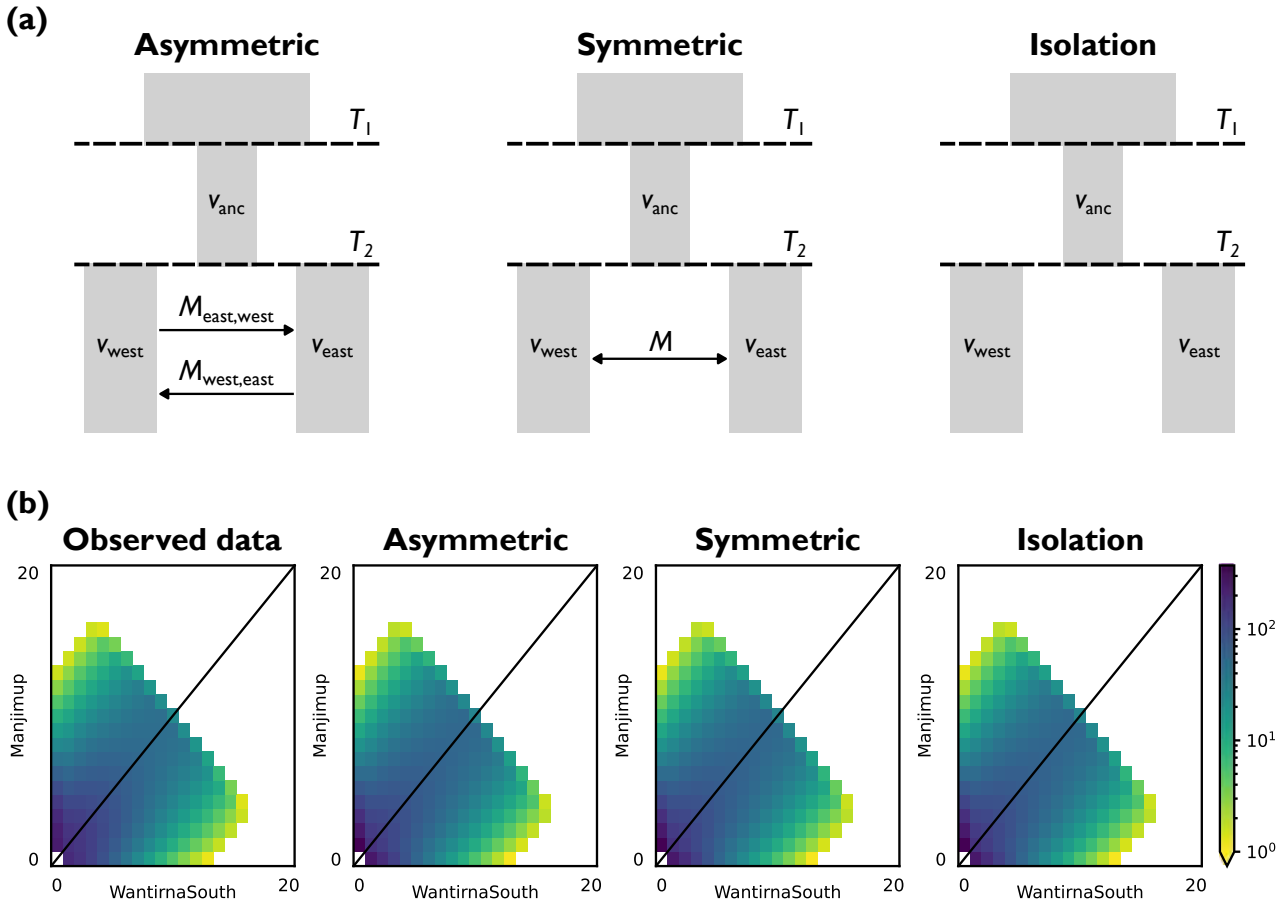

Figure 5. Demographic inference using paired eastern and western populations of *H. destructor*. (a) Schematics of the modelled demographic scenarios and parameters. (b) Representative joint site-frequency spectra (SFS) between Wantirna South (eastern population, *x*-axis) and Manjimup (western population, *y*-axis). Each plot represents the downward projected (20-by-20) joint counts for the minor allele (folded spectra). The scale bar indicates the density of joint allele counts. From left to right, the observed data, and the modelled SFS under the asymmetric migration, symmetric migration, and isolation scenarios. Modelled joint SFS and associated residual plots for all populations pairs can be found in the Supplementary Information (Figures S5–S7).

Discussion

Draft genome assembly of *Halotydeus destructor*

We report the first draft genome for the redlegged earth mite, *Halotydeus destructor*, a major pest of Australian grains and pastures. This pest is becoming increasingly difficult to manage due to ongoing evolution and spread of pesticide resistance (Arthur *et al.*, 2021). Our draft assembly represents the first genome for the mite family Penthaleidae and the eleventh genome for the order Trombidiformes. Genomic resources are under-represented in the Trombidiformes, despite 25,821 described mite species among 151 families (Zhang *et al.*, 2011). Our draft genome fills an important gap in available reference genomes. The genome of *H. destructor* is small, only 48.9 Mb in length, which constitutes one of the smallest genomes in the Trombidiformes so far (Figure 1). Of the genomes reported in NCBI, only the gall mite, *Fragariocoptes setiger* (GCA_019968535.1; 40.9 Mb), and the tomato russet mite, *Aculops lycopersici* (GCA_015350385.1; 32.5 Mb), have smaller genomes than *H. destructor*. Our draft assembly comprised 132 contigs

(scaffolds), with an N50 = 874,492 and an L50 = 14. Our longest contig was 4.59 Mb, and the smallest contig was 0.019 Mb.

Our draft genome was derived from a heterogeneous pool of *H. destructor* from a single location, Wantirna South. Despite the ease of collection, *H. destructor* is challenging to raise under laboratory conditions (Umina, 2007), limiting the production of inbred lines. The increased homozygosity in inbred lines is beneficial for genome assembly, reducing error, fragmentation, and loss of genomic regions (Rhie *et al.*, 2021; Whibley *et al.*, 2021). Our pooled sample of *H. destructor* used in genome assembly had an estimated heterozygosity rate of 1.4% (percentage of sites exhibiting polymorphism), which is comparable to those seen in other heterogeneous genomes in trombidiform mites, for example, *Leptotrombidium deliense* (1.2%) and *Dinothrombium tinctorium* (0.7%) (Dong *et al.*, 2018). Notwithstanding the heterogeneity in our sample, our draft genome for *H. destructor* is a relatively complete assembly. For instance, we recovered a major proportion of the expected single copy orthologs for arthropods (89.5%) and arachnids (90.0%). Furthermore, our assembly is on par with other genomes assembled for the Trombidiformes, in terms of the number of scaffolds relative to genome size (Figure 1). Whilst there is no information on the expected chromosome number for mites in the family Penthaleidae, karyotype studies of members of the family Tetranychidae have reported chromosomal numbers between 3 and 14 (Helle & Bolland, 1967; Feiertag-Koppen, 1980).

Amplification of *ace*, the target gene of organophosphate pesticides

Extensive amplification in the acetylecholinesterase gene, *ace*, was revealed by our genome assembly for *H. destructor* (Figures 2 & 3). *Ace* is the target gene of organophosphate pesticides, which have been used extensively in chemical control of invasive Australian *H. destructor* for decades (Arthur *et al.*, 2021). The number of *ace* copies in *H. destructor* outstripped all other select arthropod taxa in our comparative analyses, except *T. urticae*, another mite from the order Trombidiformes. *Tetranychus urticae* is the most resistant arthropod pest, evolving resistance to virtually every class of chemical used to control it (Van Leeuwen *et al.*, 2010; Sparks & Nauen, 2015). In addition to target-site mutations that inhibit interactions between *ace* and organophosphates, *ace* amplifications are also associated with increased resistance in particular strains of *T. urticae* (Khajehali *et al.*, 2010; Kwon *et al.*, 2010a; b). *Ace* amplifications may also underlie organophosphate resistance in another *Tetranychus* species, *T. evansi*, the tomato red spider mite (Carvalho *et al.*, 2012), and are prevalent in organophosphate-resistant mosquito species (Weetman *et al.*, 2018). Amplifications of pesticide target genes may be adaptive on two fronts: (1) if increased copy number increases expression, higher pesticide doses are required to suppress the activity of target genes; or (2) by maintaining different susceptible and resistant copies of a target gene, fitness costs associated with target-site mutations can be offset (Feyereisen *et al.*, 2015; Weetman *et al.*, 2018). It is likely that *ace* copy number contributes to organophosphate resistance in *H. destructor*, but further research is required to explore this thoroughly.

Chromosomal regions containing gene amplifications are difficult to assemble (Feyereisen *et al.*, 2015), with the risk that homologous genomic regions are split into separate contigs due to excessive sequence heterogeneity. However, it is unlikely that our assembly has suffered much from contig over-splitting. There was minimal duplication of both arthropod (1.4%) and arachnid (2.4%) single copy orthologs in our *H. destructor* assembly. We also observed a single copy of elongation factor 1-α (HDE_08018), a common standard for copy number analyses, and a single copy of the *para* sodium channel (HDE_05193), which is the target gene for pyrethroid pesticides. Nonetheless, there is at least one example where genomic structural variation around amplified *ace* genes have led to bioinformatic separation into different contigs. In our neighbour joining trees of inferred *ace* protein sequences, *ace* genes from contigs 344 and contig 1156 exhibited suspect across-contig relationships (Figure 3; but see also Figure S2). *Ace g*enes HDE_06662 and HDE_06664 on contig 344 were both sister to HDE_00765 on contig 1156, whereas HDE_06661 on contig 344 was sister to HDE_00766 on contig 1156. HDE_06662 and HDE_06664 both differed from HDE_00765 by a total of 66

residues, and HDE_06661 differed from HDE_00766 by just four residues. These groups of related *ace* genes formed a monophyletic group with HDE_06666, which is also on contig 344. Closer examination revealed that contig 1156 (29.53 Kb) comprises a truncated and inverted region of contig 344 (145.58 Kb): contig 1156 starts at the tail-end of contig 344's HDE_06662 *ace* gene, extending backwards (Figure S4).

Not only does the *H. destructor* genome comprise different copies of the *ace* gene, but some of these copies may be associated with structural rearrangements at different genomic scales. Any role of structural variants in pesticide resistance adaptation in *H. destructor* remains to be determined, although chromosomal rearrangements can be important for bringing together genes into adaptive complexes and blocking recombination that might otherwise dissociate these adaptive complexes (Charlesworth & Charlesworth, 1973; Hoffmann *et al.*, 2004; Li *et al.*, 2022). Analysis of SNP variation within the *ace* genes indicates local populations harbour unique variants of the different *ace* gene copies (Figure 4e). The issue of whether similar selection pressures result in different genetic outcomes has been an enduring interest in evolutionary biology (Cohan & Hoffmann, 1989; Arendt & Reznick, 2008; Barghi *et al.*, 2020), with recent genomic and transcriptomic work highlighting both unique responses of related populations to similar selection pressures as well as divergent responses when polygenic changes are expected (Griffin *et al.*, 2017; Seabra *et al.*, 2018; Sprengelmeyer & Pool, 2021). Localised population-specific genetic changes may prove challenging when attempting to develop predictive tools for evolutionary changes, such as diagnostic molecular markers, for screening resistance (Thia *et al.*, 2021a). Currently, we are further investigating the potentially complex interplay between *ace* copy number, *ace* alleles, their associated structural arrangements, and additional genomic factors on organophosphate resistance in *H. destructor*.

Regional patterns of genetic structure and diversity

Our population genomic analyses revealed genetic differences among populations of *H. destructor* within and between eastern and western regions of Australia (Figure 4). Among our sampled populations, genetic differentiation was greater between regions (0.032 ≤ pairwise F_{ST} ≤ 0.068) relative to within regions (0.014 ≤ pairwise F_{ST} ≤ 0.039). Moreover, western populations had significantly higher genomic heterozygosity (Figure 4b) and larger inferred population sizes (Table 1) relative to eastern populations. These observations are consistent with the stepping-stone invasion pathway of *H. destructor* into Australia, whereby the first incursion was in Western Australian, which acted as the source for invasions into eastern regions (Ridsdill-Smith, 1997).

The clear regional structuring of genetic variation suggests that gene flow does not homogenise Australian populations over large geographic distances. However, our demographic analyses were equivocal on the role of gene flow in shaping patterns of genetic variation. The best model of the joint SFS between eastern and western *H. destructor* populations was always the isolation scenario (Table 1). However, the best replicate simulations for the symmetric and asymmetric migration scenarios had virtually identical fits to the observed data (Figures 5 & S5–S7) with ΔAIC = –2 and –4, respectively. *Halotydeus destructor* invaded Western Australia around 1917, and eastern Australia around 1921 (Ridsdill-Smith, 1997). This short evolutionary history likely poses a challenge for demographic inference; insufficient time makes it difficult to distinguish between scenarios of recent divergence with isolation and ongoing gene flow. Indeed, many values of gene flow produced comparable fits to the observed SFS between eastern and western *H. destructor* populations (Figures S8, S9, S11, & S12). Although a scenario of recent divergence with isolation is the most parsimonious, migration may help maintain some level of genetic connection between regional populations.

Quantifying the magnitude of gene flow among populations has been a long-standing question in the management of *H. destructor* (Weeks *et al.*, 1995; Qin, 1997). Understanding the rates of dispersal between regions is necessary for predicting the risk of pesticide resistance alleles moving between states. Yang *et al.* (2020) showed that Australian-wide pyrethroid resistance has often evolved through parallel evolution in local populations, but also less frequently via

spread of dispersing resistance alleles. This conclusion was based on contrasting genetic differentiation among resistant and susceptible populations at different spatial scales. Yang *et al.* (2020) also found that *H. destructor* can exhibit isolation-by-distance over very short spatial distances (100s of meters), but this does not exclude the possibility of occasional long distance dispersal, such as that required to rapidly spread this pest across Australia within a decade of its introduction (Ridsdill-Smith, 1997). In our demographic scenarios, migration was either absent or constant through time. However, the dynamics of *H. destructor* dispersal is expected to be variable; on average, dispersal may be local and spatially constricted, but probably coupled with rare long-distance dispersal events mediated by wind and (or) the movement of livestock, fodder and machinery. Such a hypothesis might be tested using individual-level simulations and probabilistic dispersal kernels to compare the effects of ongoing migration with the rare dispersal of propagules between regions. Our ability to infer demographic parameters was limited by our pool-seq design; whole genome sequences of individuals provide a useful target for future work because they can tap into rich linkage disequilibrium and haplotype information, providing greater power to discern between competing models of gene flow (Sousa & Hey, 2013; Smith & Flaxman, 2020; Marchi *et al.*, 2021).

Conclusions

We have developed a high-quality draft genome of the mite pest, *H. destructor*. We show that *H. destructor* exhibits a large radiation of *ace* genes reminiscent of the radiation detected in spider mites, which may be linked to selection by organophosphate pesticides. Patterns of genetic variation and estimates of relative effective population sizes are congruent with serial invasions of *H. destructor* from west to east. We examined the joint SFS between eastern and western populations to make demographic inferences, but these were limited by the relatively short evolutionary timescale *H. destructor* has been in its invasive range. Our draft genome for *H. destructor*, in combination with the recently published mitogenome (Thia *et al.*, 2021b), adds to the depauperate genomic resources available for mites in the order Trombidiformes. These genomic resources will facilitate deeper investigations into the ecology and evolution of this important invasive pest in Australian grains.

Acknowledgements

This work was funded by the Grains Research Development Corporation (Australia) as part of the Australian Grains Pest Innovation Program, and with support from the University of Melbourne. Computational resources were provided by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative.

Data accessibility

Our draft genome for *H. destructor* is available on GenBank (Assembly, ASM2275052v1; Genome, GCA_022750525.1). All sequencing data has been submitted under the GenBank Project PRJNA756307. Submissions to the GenBank Sequence Read Archive include: Illumina short-reads for genome assembly (Wantirna South population, SRR15698717); PacBio long-reads for genome assembly (Wantirna South population, SRR15698716), Illumina short-reads for transcriptomes (Capel and Manjimup populations, being processed), and Illumina short-reads for population genomics (see Supporting Information, being processed).

Author Contributions

JT, PU, OE, TW and AH contributed to the design of this project. JT, OE and PU collected mite populations. OW and TW generated the transcriptome data. JT and QY collected the nuclear genome data. JT, PK, NY and RG contributed to the assembly of the genome. JT led the comparative analyses, population genomics, and demographic modelling. JT wrote the original manuscript. All authors contributed toward revisions of the manuscript.

References

Arendt, J. & Reznick, D. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**: 26–32.

Arnold, B., Corbett-Detig, R.B., Hartl, D. & Bomblies, K. 2013. RAD seq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Mol. Ecol.* **22**: 3179–3190.

Arthur, A.L., Maino, J., Hoffmann, A.A., Jasper, M., Lord, A., Micic, S., *et al.* 2021. Learnings from over a decade of increasing pesticide resistance in the redlegged earth mite, *Halotydeus destructor* (Tucker). *Pest Manag. Sci.* **77**: 3013–3024.

Barghi, N., Hermisson, J. & Schlötterer, C. 2020. Polygenic adaptation: a unifying framework to understand positive selection. *Nat. Rev. Genet.* **21**: 769–781. Springer US.

Bass, C., Denholm, I., Williamson, M.S. & Nauen, R. 2015. The global status of insect resistance to neonicotinoid insecticides. *Pestic. Biochem. Physiol.* **121**: 78–87.

Bates, D., Mächler, M., Bolker, B. & Walker, S. 2014. Fitting linear mixed-effects models using lme4. *arXiv Prepr.*

Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C. & Hochreiter, S. 2015. msa: an R package for multiple sequence alignment. *Bioinformatics* **31**: 3997–3999.

Burnham, K.P. & Anderson, D.R. 2004. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**: 261–304. Sage Publications Sage CA: Thousand Oaks, CA.

Cao, L.J., Song, W., Chen, J.C., Fan, X.L., Hoffmann, A.A. & Wei, S.J. 2022. Population genomic signatures of the oriental fruit moth related to the Pleistocene climates. *Commun. Biol.* **5**: 1–12. Springer US.

Carvalho, R., Yang, Y., Field, L.M., Gorman, K., Moores, G., Williamson, M.S., *et al.* 2012. Chlorpyrifos resistance is associated with mutation and amplification of the acetylcholinesterase-1 gene in the tomato red spider mite, *Tetranychus evansi*. *Pestic. Biochem. Physiol.* **104**: 143–149. Elsevier Inc.

Charlesworth, B. & Charlesworth, D. 1973. Selection of new inversions in multi-locus genetic systems. *Genet. Res.* **21**: 167–183.

Chen, L., Sun, J.T., Jin, P.Y., Hoffmann, A.A., Bing, X.L., Zhao, D.S., *et al.* 2020. Population genomic data in spider mites point to a role for local adaptation in shaping range shifts. *Evol. Appl.* **13**: 2821–2835.

Chen, S., Zhou, Y., Chen, Y. & Gu, J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890.

Cheng, X., Umina, P.A., Lee, S.F. & Hoffmann, A.A. 2019. Pyrethroid resistance in the pest mite, *Halotydeus destructor*: Dominance patterns and a new method for resistance screening. *Pestic. Biochem. Physiol.* **159**: 9–16.

Cohan, F.M. & Hoffmann, A.A. 1989. Uniform selection as a diversifying force in evolution: evidence from Drosophila.

Am. Nat. 134: 613–637.

Conesa, A. & Götz, S. 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* 2008. Hindawi.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. & Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674–3676. Oxford University Press.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., *et al.* 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10: giab008. Oxford University Press.

Devonshire, A.L. & Moores, G.D. 1982. A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (*Myzus persicae*). *Pestic. Biochem. Physiol.* 18: 235–246.

Dong, K., Du, Y., Rinkevich, F., Nomura, Y., Xu, P., Wang, L., *et al.* 2014. Molecular biology of insect sodium channels and pyrethroid resistance. *Insect Biochem. Mol. Biol.* 50: 1–17.

Dong, X., Chaisiri, K., Xia, D., Armstrong, S.D., Fang, Y., Donnelly, M.J., *et al.* 2018. Genomes of trombidid mites reveal novel predicted allergens and laterally transferred genes associated with secondary metabolism. *Gigascience* 7: 1–33. Oxford University Press.

Douris, V., Steinbach, D., Panteleri, R., Livadaras, I., Pickett, J.A., Van Leeuwen, T., *et al.* 2016. Resistance mutation conserved between insects and mites unravels the benzoylurea insecticide mode of action on chitin biosynthesis. *PNAS* 113: 14692–14697.

Edwards, O.R., Walsh, T.K., Metcalfe, S., Tay, W.T., Hoffmann, A.A., Mangano, P., *et al.* 2018. A genomic approach to identify and monitor a novel pyrethroid resistance mutation in the redlegged earth mite, *Halotydeus destructor*. *Pestic. Biochem. Physiol.* 144: 83–90.

Feiertag-Koppen, C.C.M. 1980. Cytological studies of the two-spotted spider mite *Tetranychus urticae* Koch (tetranychidae, trombidiformes) II. Meiosis in growing oocytes. *Genetica* 54: 173–180.

Ferretti, L., Ramos-Onsins, S.E. & Pérez-Enciso, M. 2013. Population genomics from pool sequencing. *Mol. Ecol.* 22: 5561–5576.

Feyereisen, R., Dermauw, W. & Van Leeuwen, T. 2015. Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods. *Pestic. Biochem. Physiol.* 121: 61–77.

Garrison, E. & Marth, G. 2012. Haplotype-based variant detection from short-read sequencing. *arXiv* 1207.3907.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., *et al.* 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36: 3420–3435. Oxford University Press.

Graves, S., Piepho, H.-P. & Selzer, L. 2019. multcompView: Visualizations of Paired Comparisons.

Griffin, P.C., Hangartner, S.B., Fournier-Level, A. & Hoffmann, A.A. 2017. Genomic trajectories to desiccation resistance: Convergence and divergence among replicate selected *Drosophila* lines. *Genetics* 205: 871–890.

Gunning, R. V., Easton, C.S., Balfe, M.E. & Ferris, I.G. 1991. Pyrethroid resistance mechanisms in Australian *Helicoverpa armigera*. *Pestic. Sci.* 33: 473–490.

Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H. & Bustamante, C.D. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**: e1000695.

Helle, W. & Bolland, H.R. 1967. Karyotypes and sex-determination in spider mites (Tetranychidae). *Genetica* **38**: 43–53.

Hivert, V., Leblois, R., Petit, E.J., Gautier, M. & Vitalis, R. 2018. Measuring genetic differentiation from pool-seq data. *Genetics* **210**: 315–330.

Hoffmann, A.A., Porter, S. & Kovacs, I. 1997. The response of the major crop and pasture pest, the red-legged earth mite (*Halotydeus destructor*) to pesticides: Dose-response curves and evidence for tolerance. *Exp. Appl. Acarol.* **21**: 151–162.

Hoffmann, A.A., Sgrò, C.M. & Weeks, A.R. 2004. Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.* **19**: 482–488.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**: D109–D114.

Khajehali, J., van Leeuwen, T., Grispou, M., Morou, E., Alout, H., Weill, M., *et al.* 2010. Acetylcholinesterase point mutations in European strains of *Tetranychus urticae* (Acari: Tetranychidae) resistant to organophosphates. *Pest Manag. Sci.* **66**: 220–228.

Korhonen, P.K., Hall, R.S., Young, N.D. & Gasser, R.B. 2019. Common workflow language (CWL)-based software pipeline for de novo genome assembly from long-and short-read data. *Gigascience* **8**: giz014.

Korhonen, P.K., Kinkar, L., Young, N.D., Cai, H., Lightowlers, M.W., Gauci, C., *et al.* 2022. Chromosome-scale Echinococcus granulosus (genotype G1) genome reveals the Eg95 gene family and conservation of the EG95-vaccine molecule. *Commun. Biol.* **5**: 1–10.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**: 1–26.

Kwon, D.H., Clarkt, J.M. & Lee, S.H. 2010a. Extensive gene duplication of acetylcholinesterase associated with organophosphate resistance in the two-spotted spider mite. *Insect Mol. Biol.* **19**: 195–204.

Kwon, D.H., Im, J.S., Ahn, J.J., Lee, J.H., Marshall Clark, J. & Lee, S.H. 2010b. Acetylcholinesterase point mutations putatively associated with monocrotophos resistance in the two-spotted spider mite. *Pestic. Biochem. Physiol.* **96**: 36–42.

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993. Oxford University Press.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li, Q., Lindtke, D., Rodríguez-Ramírez, C., Kakioka, R., Takahashi, H., Toyoda, A., *et al.* 2022. Local adaptation and the evolution of genome architecture in threespine stickleback. *Genome Biol. Evol.*

Liu, H., Wu, S., Li, A. & Ruan, J. 2021. SMARTdenovo: a *de novo* assembler using long noisy reads. *Gigabyte* **2021**: 1–9. GigaScience Press.

Magrane, M. 2011. UniProt Knowledgebase: a hub of integrated protein data. *Database* **2011**.

Maino, J.L., Umina, P.A. & Hoffmann, A.A. 2018. Climate contributes to the evolution of pesticide resistance. *Glob. Ecol. Biogeogr.* **27**: 223–232.

Marcais, G. & Kingsford, C. 2012. Jellyfish: A fast k-mer counter. *Tutorialis e Manuais* **1**: 1–8.

Marchi, N., Schlichta, F. & Excoffier, L. 2021. Demographic inference. *Curr. Biol.* **31**: R276–R279.

Newman, L.J. 1925. The red legged earth mite *Penthaleus destructor* (Jack). *J. Dep. Agric. West. Aust.* **2**: 469–475.

Pagès, H., Aboyoun, P., Gentleman, R. & DebRoy, S. 2017. Biostrings: Efficient manipulation of biological strings.

Paradis, E., Claude, J. & Strimmer, K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290. Oxford University Press.

Pruitt, K.D., Tatusova, T., Brown, G.R. & Maglott, D.R. 2012. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**: D130–D135.

Puinean, A.M., Foster, S.P., Oliphant, L., Denholm, I., Field, L.M., Millar, N.S., *et al.* 2010. Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae*. *PLoS Genet.* **6**: 1–11.

Qin, T.-K. 1997. Population genetics of redlegged earth mites *Halotydeus destructor* and *H. anthropus* (Acarina: Penthaleidae) from Australia and/or South Africa. *Bull. Entomol. Res.* **87**: 289–298. University of Melbourne Library.

Rhie, A., McCarthy, S.A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., *et al.* 2021. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**: 737–746.

Ridsdill-Smith, T.J. 1997. Biology and control of *Halotydeus destructor* (Tucker) (Acarina: Penthaleidae): A review. *Exp. Appl. Acarol.* **21**: 195–224.

Rispe, C., Legeai, F., Nabity, P.D., Fernández, R., Arora, A.K., Baa-Puyoulet, P., *et al.* 2020. The genome sequence of the grape phylloxera provides insights into the evolution, adaptation, and invasion routes of an iconic pest. *BMC Biol.* **18**: 1–25. BMC Biology.

Robinson, M.T. & Hoffmann, A.A. 2000. Additional tests on the effects of pesticides on cryptic species of blue oat mite (*Penthaleus* spp.) and the redlegged earth mite (*Halotydeus destructor*). *Aust. J. Exp. Agric.* **40**: 671–678.

Rodríguez-Ezpeleta, N., Bradbury, I.R., Mendibil, I., Álvarez, P., Cotano, U. & Irigoien, X. 2016. Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: Effects of sequence clustering parameters and hierarchical SNP selection. *Mol. Ecol. Resour.* **16**: 991–1001.

Russell, R.J., Claudianos, C., Campbell, P.M., Horne, I., Sutherland, T.D. & Oakeshott, J.G. 2004. Two major classes of target site insensitivity mutations confer resistance to organophosphate and carbamate insecticides. *Pestic. Biochem. Physiol.* **79**: 84–93.

Seabra, S.G., Fragata, I., Antunes, M.A., Faria, G.S., Santos, M.A., Sousa, V.C., *et al.* 2018. Different genomic changes underlie adaptive evolution in populations of contrasting history. *Mol. Biol. Evol.* **35**: 549–563.

Shafer, A., Peart, C.R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C.W., *et al.* 2016. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol. Evol.*

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

Smith, C.C.R. & Flaxman, S.M. 2020. Leveraging whole genome sequencing data for demographic inference with approximate Bayesian computation. *Mol. Ecol. Resour.* **20**: 125–139.

Sousa, V. & Hey, J. 2013. Understanding the origin of species with genome-scale data: Modelling gene flow. *Nat. Rev.*

Genet. **14**: 404–414.

Sparks, T.C., Lockwood, J.A., Byford, R.L., Graves, J.B. & Leonard, B.R. 1989. The role of behavior in insecticide resistance. *Pestic. Sci.* **26**: 383–399.

Sparks, T.C. & Nauen, R. 2015. IRAC: Mode of action classification and insecticide resistance management. *Pestic. Biochem. Physiol.* **121**: 122–128.

Sprengelmeyer, Q.D. & Pool, J.E. 2021. Ethanol resistance in Drosophila melanogaster has increased in parallel cold-adapted populations and shows a variable genetic architecture within and between populations. *Ecol. Evol.* **11**: 15364–15376.

Tay, W.T. & Gordon, K.H.J. 2019. Going global – genomic insights into insect invasions. *Curr. Opin. Insect Sci.* **31**: 123–130. Elsevier Inc.

Tay, W.T., Rane, R. V., Padovan, A., Walsh, T.K., Elfekih, S., Downes, S., *et al.* 2022. Global population genomic signature of *Spodoptera frugiperda* (fall armyworm) supports complex introduction events across the Old World. *Commun. Biol.* **5**: 1–15. Springer US.

Thia, J.A., Hoffmann, A.A. & Umina, P.A. 2021a. Empowering Australian insecticide resistance research with genetic information: the road ahead. *Austral Entomol.* **60**: 147–162.

Thia, J.A. & Riginos, C. 2019. *genomalicious*: serving up a smorgasbord of R functions for population genomic analyses. *bioRxiv* 667337.

Thia, J.A., Young, N.D., Korhnen, P.K., Yang, Q., Gasser, R.B., Umina, P.A., *et al.* 2021b. The mitogenome of Halotydeus destructor (Tucker) and its relationships with other trombidiform mites as inferred from nucleotide sequences and gene arrangements. *Ecol. Evol.* **11**: 14162–14174.

Umina, P.A. 2007. Pyrethroid resistance discovered in a major agricultural pest in southern Australia: the redlegged earth mite *Halotydeus destructor* (Acari: Penthaleidae). *Pest Manag. Sci.* **63**: 1185–1190.

Umina, P.A., Lord, A., Micic, S. & Edwards, O. 2017. Discovery and characterisation of field resistance to organophosphorus chemicals in a major mite pest, *Halotydeus destructor*. *Pest Manag. Sci.* **73**: 1719–1724.

Umina, P.A., McDonald, G., Maino, J., Edwards, O. & Hoffmann, A.A. 2019. Escalating insecticide resistance in Australian grain pests: contributing factors, industry trends and management opportunities. *Pest Manag. Sci.* **75**: 1494–1506.

Van Leeuwen, T., Vontas, J., Tsagkarakou, A., Dermauw, W. & Tirry, L. 2010. Acaricide resistance mechanisms in the two-spotted spider mite *Tetranychus urticae* and other important Acari: A review. *Insect Biochem. Mol. Biol.* **40**: 563–572.

Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., *et al.* 2017. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**: 2202–2204.

Walsh, T.K., Joussen, N., Tian, K., Mcgaughran, A., Anderson, C.J., Qiu, X., *et al.* 2018. Multiple recombination events between two cytochrome P450 loci contribute to global pyrethroid resistance in *Helicoverpa armigera*. *PLoS One* **13**: e0197760.

Weeks, A.R., Fripp, Y.J. & Hoffmann, A.A. 1995. Genetic structure of *Halotydeus destructor* and *Penthaleus major* populations in Victoria (Acari: Penthaleidae). *Exp. Appl. Acarol.* **19**: 633–646.

Weetman, D., Djogbenou, L.S. & Lucas, E. 2018. Copy number variation (CNV) and insecticide resistance in mosquitoes: evolving knowledge or an evolving problem? *Curr. Opin. Insect Sci.* **27**: 82–88. Elsevier Inc.

Weir, B.S. & Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution (N. Y).* **38**: 1358–1370.

Whibley, A., Kelley, J.L. & Narum, S.R. 2021. The changing face of genome assemblies: Guidance on achieving high-quality reference genomes. *Mol. Ecol. Resour.* **21**: 641–652.

Yang, Q., Umina, P.A., Rašić, G., Bell, N., Fang, J., Lord, A., *et al.* 2020. Origin of resistance to pyrethroids in the redlegged earth mite (*Halotydeus destructor*) in Australia: repeated local evolution and migration. *Pest Manag. Sci.*, doi: 10.1002/ps.5538.

Zago, H.B., Siqueira, H.A.A., Pereira, E.J.G., Picanço, M.C. & Barros, R. 2014. Resistance and behavioural response of *Plutella xylostella* (Lepidoptera: Plutellidae) populations to *Bacillus thuringiensis* formulations. *Pest Manag. Sci.* **70**: 488–495.

Zdobnov, E.M. & Apweiler, R. 2001. InterProScan–an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**: 847–848.

Zhang, L., Liu, B., Zhen, W., Liu, C., Zhang, D., Zhao, S., *et al.* 2020. Genetic structure and insecticide resistance characteristics of fall armyworm populations invading China. *Mol. Ecol.* 1–15.

Zhang, Z.-Q., Fan, Q.-H., Pesic, V., Smit, H., Bochkov, A. V, Khaustov, A.A., *et al.* 2011. Order Trombidiformes Reuter, 1909. In: Zhang, Z.-Q.(Ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* **3148**: 129–138.

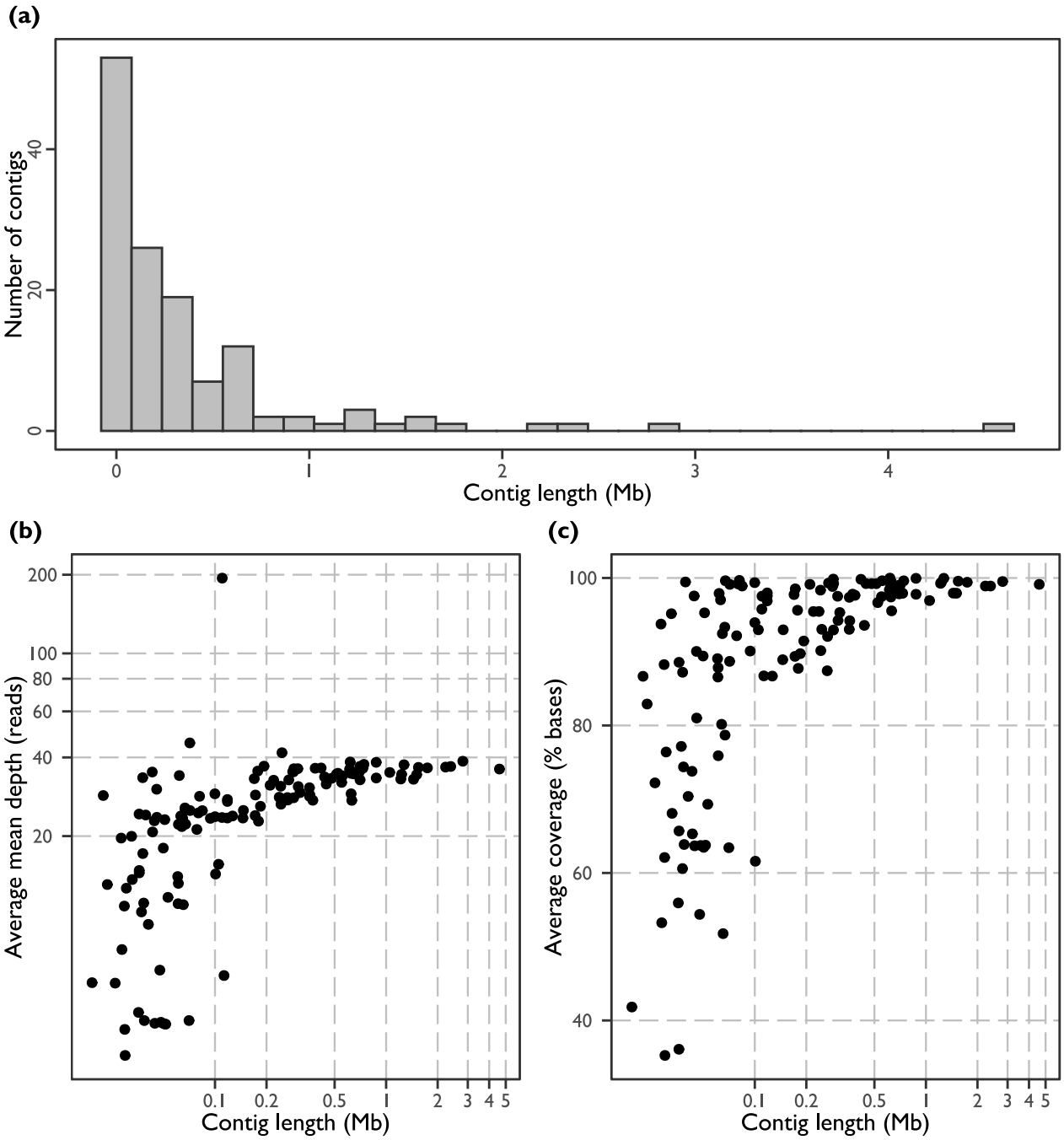

Figure S1. Contig size distributions and mapping efficiency. (a) Distribution of contig sizes in the *Halotydeus destructor* draft genome. (b) Mapping depth across the genome. The *x*-axis is the contig length and the *y*-axis is the average across sample mean depth. That is, for each contig, the mean depth was calculated for each sample, then the average was taken across samples. (c) Average coverage across the genome. The *x*-axis is the contig length and the *y*-axis is the percentage of bases covered.

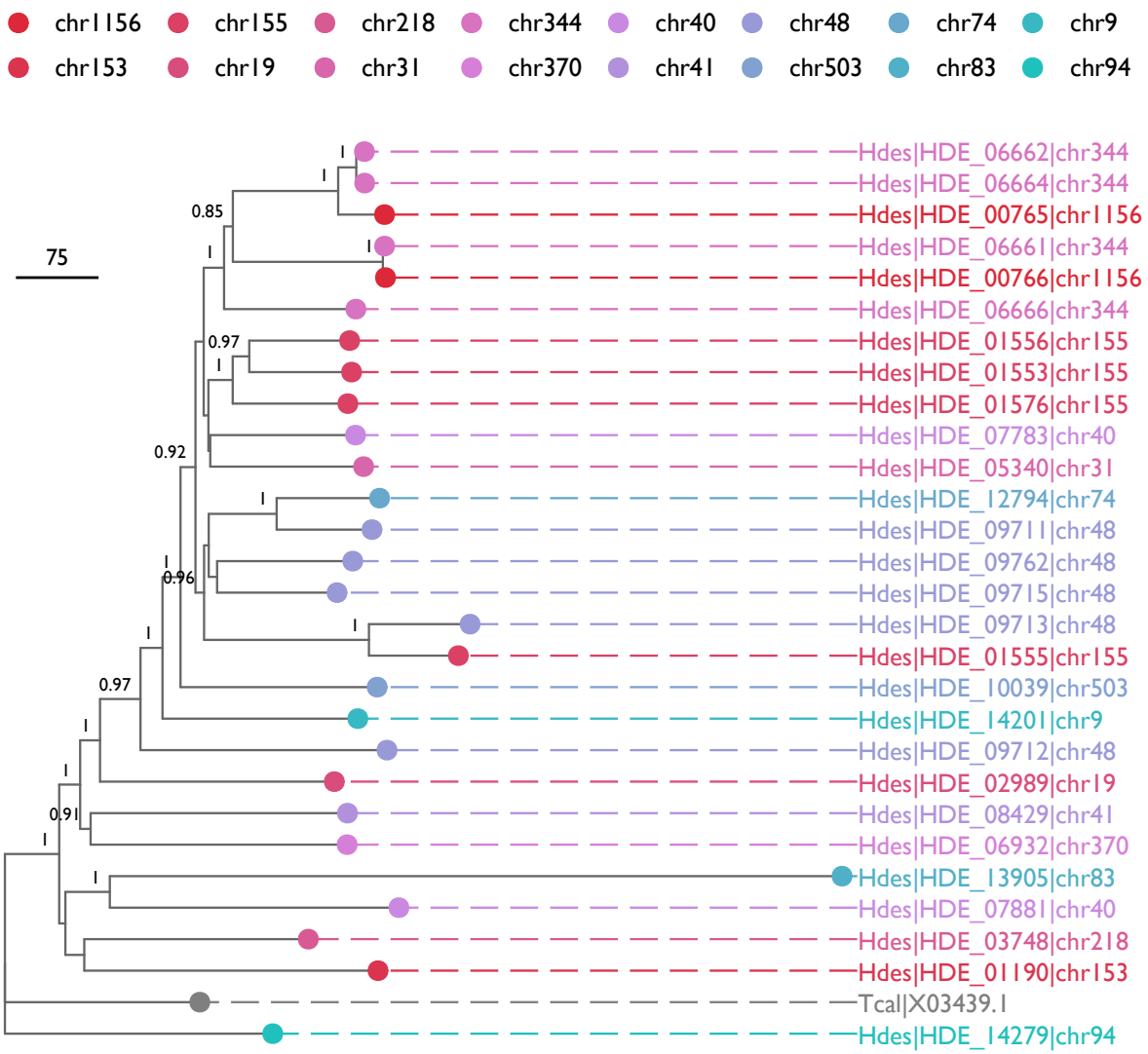

Figure S2. Neighbour joining tree of *H. destructor* ace protein sequences ("Hdes"). Each tip represents an annotated *ace* gene, with points and tip labels coloured by contig (see legend). The *ace* sequence from *Torpedo californica* ("Tcal") was used to root the tree. The scale bar indicates the number of amino acid differences. Nodes labels indicate ≥80% bootstrap support.

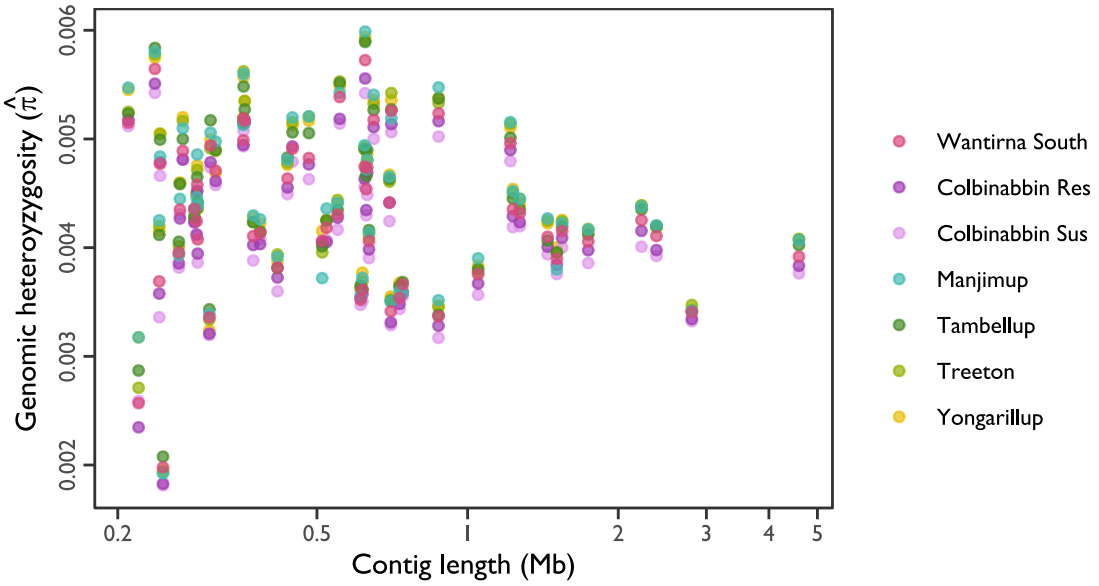

Figure S3. Genomic heterozygosity ($\hat{\pi}$) as a function of contig length. The *x*-axis is the contig length and the *y*-axis is the mean $\hat{\pi}$ at the contig level. Coloured points represent populations (see legend).

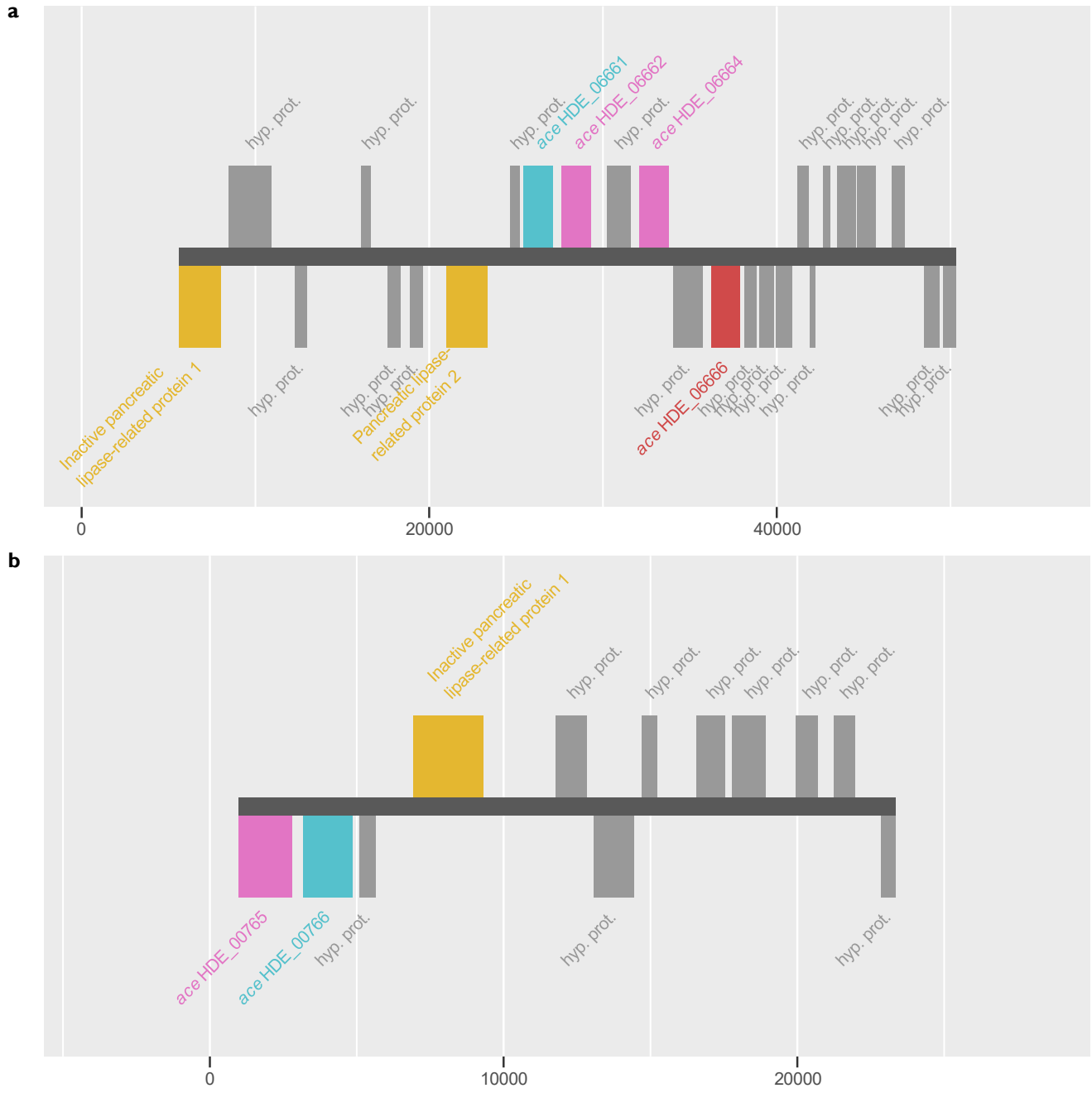

Figure S4. Putative chromosomal inversions of *H. destructor* that contain *ace* genes. The *x*-axis indicates the sequence position (bp) along a genomic contig. The dark grey horizontal bar in the center of the plot represents a subset region of a genomic contig, 20,000 bp either side of focal *ace* genes: (a) contig 344, (b) contig 1156. Blocks situated on the genomic contigs represent annotated genes. Light grey blocks represent hypothetical proteins, "hyp. prot.". Gold blocks represent two pancreatic lipase-related-like proteins. Blue, pink, and red blocks represent *ace* genes: blue represents the HDE_06661–HDE_00766 sister pair, pink represents the HDE_06662/HDE_06664–HDE_00766 sister pair; and red represents HDE_06666, which is sister to the HDE_06661–HDE_00766 and HDE_06662/HDE_06664–HDE_00766 pairs (see Figure S2).

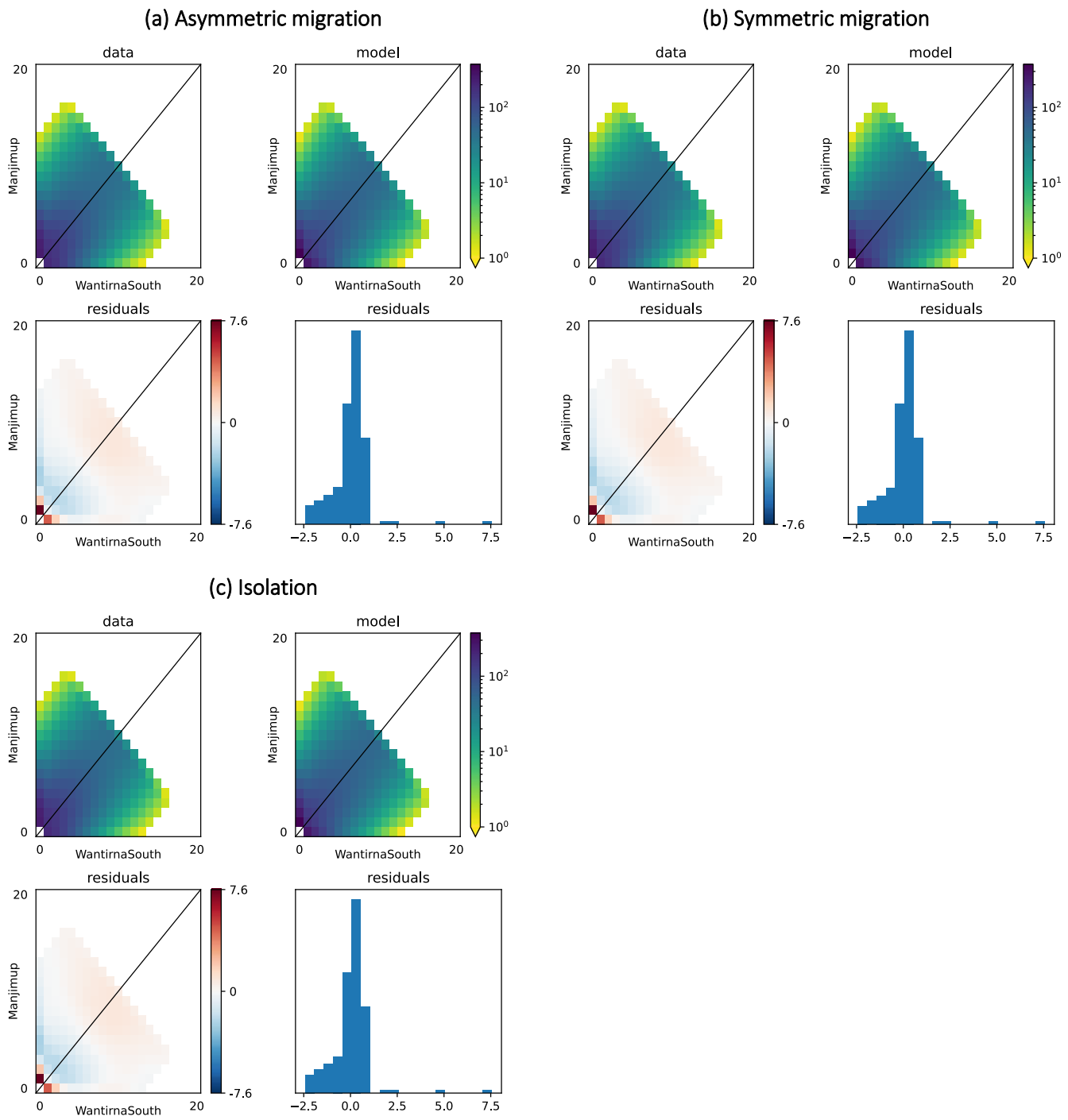

Figure S5. Observed and modelled site frequency spectra (SFS) and residual plots for the best replicate simulation, for each modelled scenario, for the Manjimup and Wantirna South population pair. In each subfigure: (top left) The observed SFS; (top right) The modelled SFS; (bottom left) The residuals mapped to the SFS grid; (bottom right) Histogram of residuals. Residuals are the Anscombe residuals, as calculated in *dadi*, effectively depicting the modelled counts minus the observed counts.

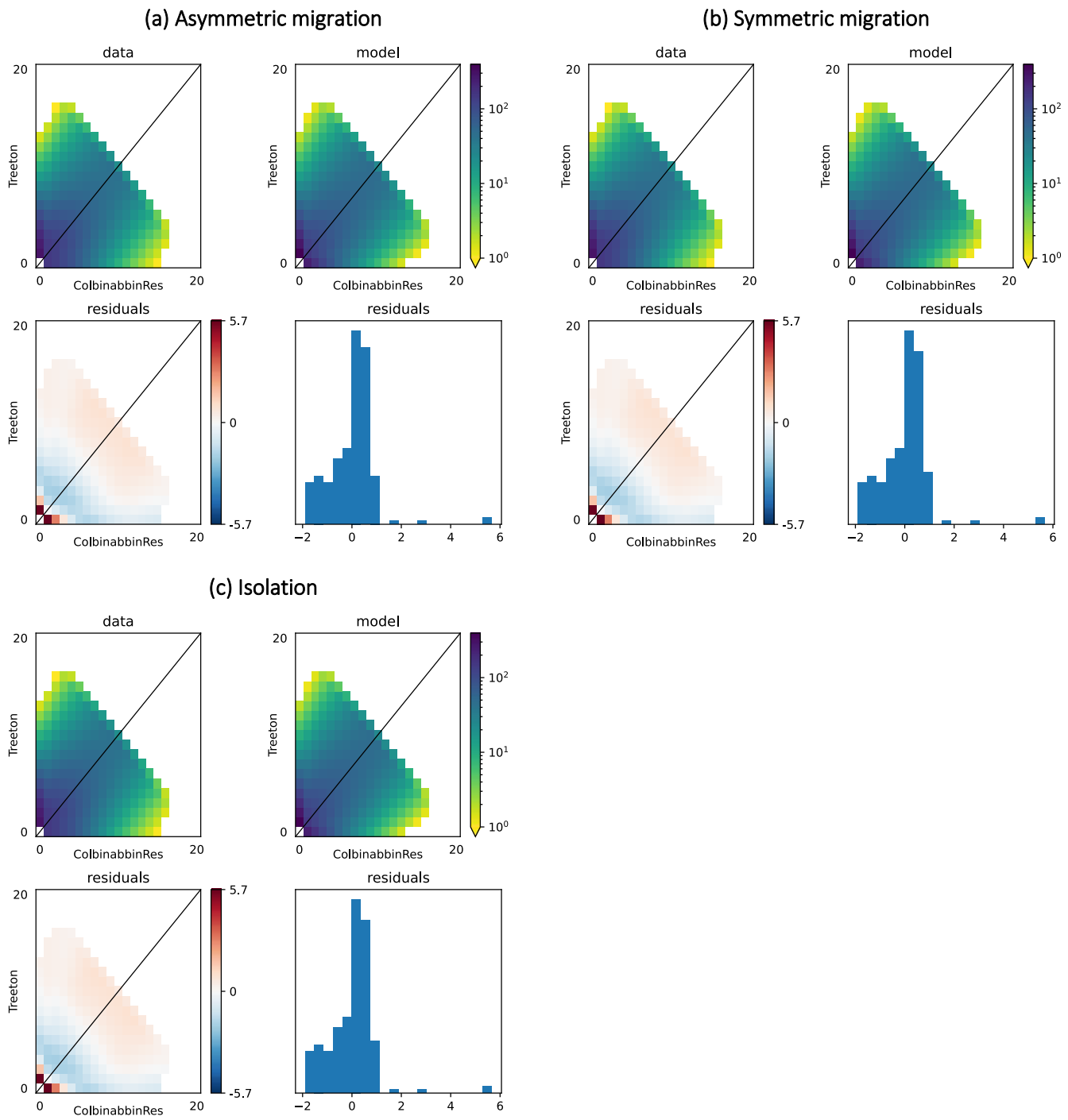

Figure S6. Observed and modelled site frequency spectra (SFS) and residual plots for the best replicate simulation, for each modelled scenario, for the Treeton and Colbinabbin Res population pair. In each subfigure: (top left) The observed SFS; (top right) The modelled SFS; (bottom left) The residuals mapped to the SFS grid; (bottom right) Histogram of residuals. Residuals are the Anscombe residuals, as calculated in *dadi*, effectively depicting the modelled counts minus the observed counts.

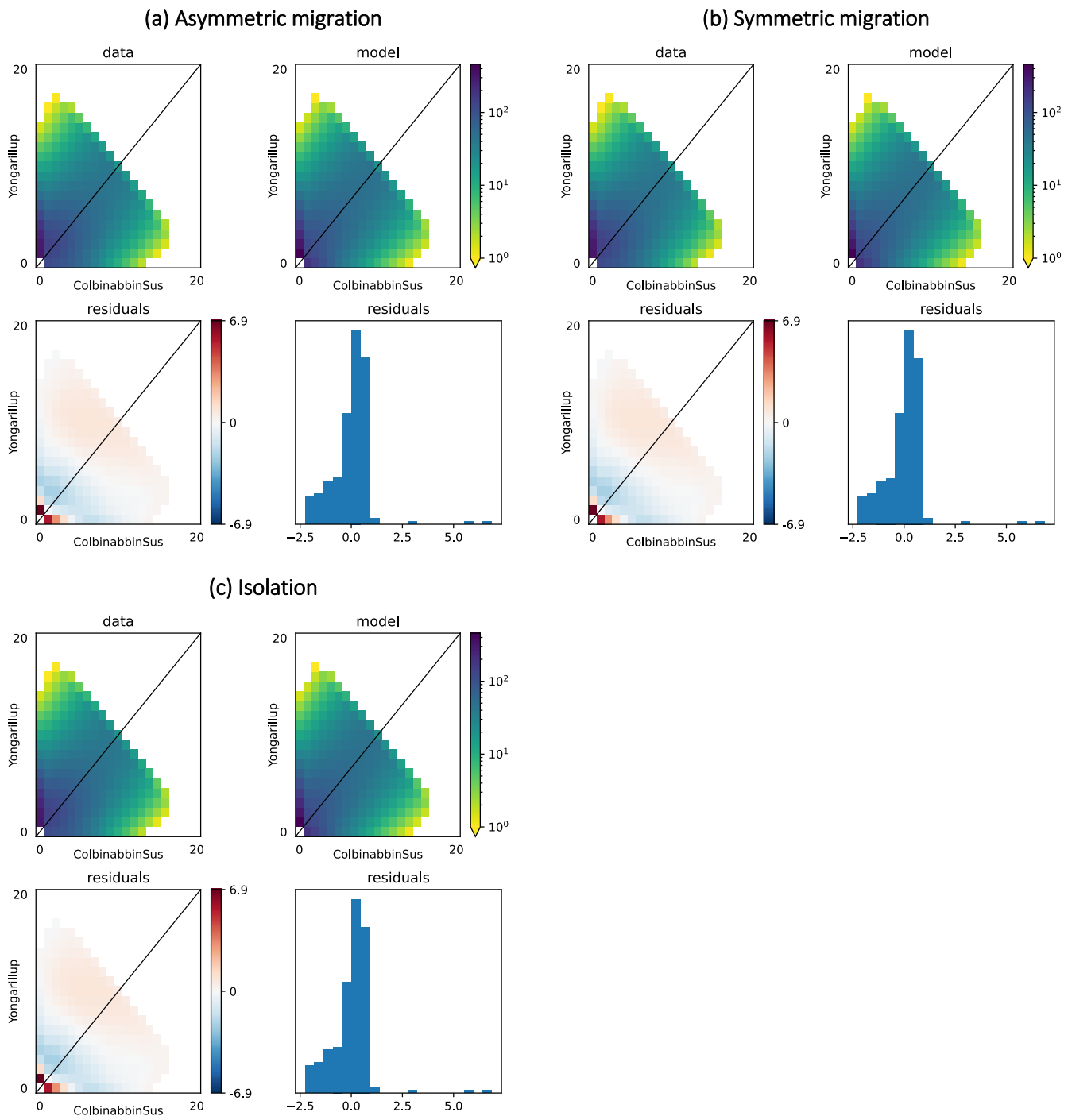

Figure S7. Observed and modelled site frequency spectra (SFS) and residual plots for the best replicate simulation, for each modelled scenario, for the Yongarillup and Colbinabbin Sus population pair. In each subfigure: (top left) The observed SFS; (top right) The modelled SFS; (bottom left) The residuals mapped to the SFS grid; (bottom right) Histogram of residuals. Residuals are the Anscombe residuals, as calculated in *dadi*, effectively depicting the modelled counts minus the observed counts.

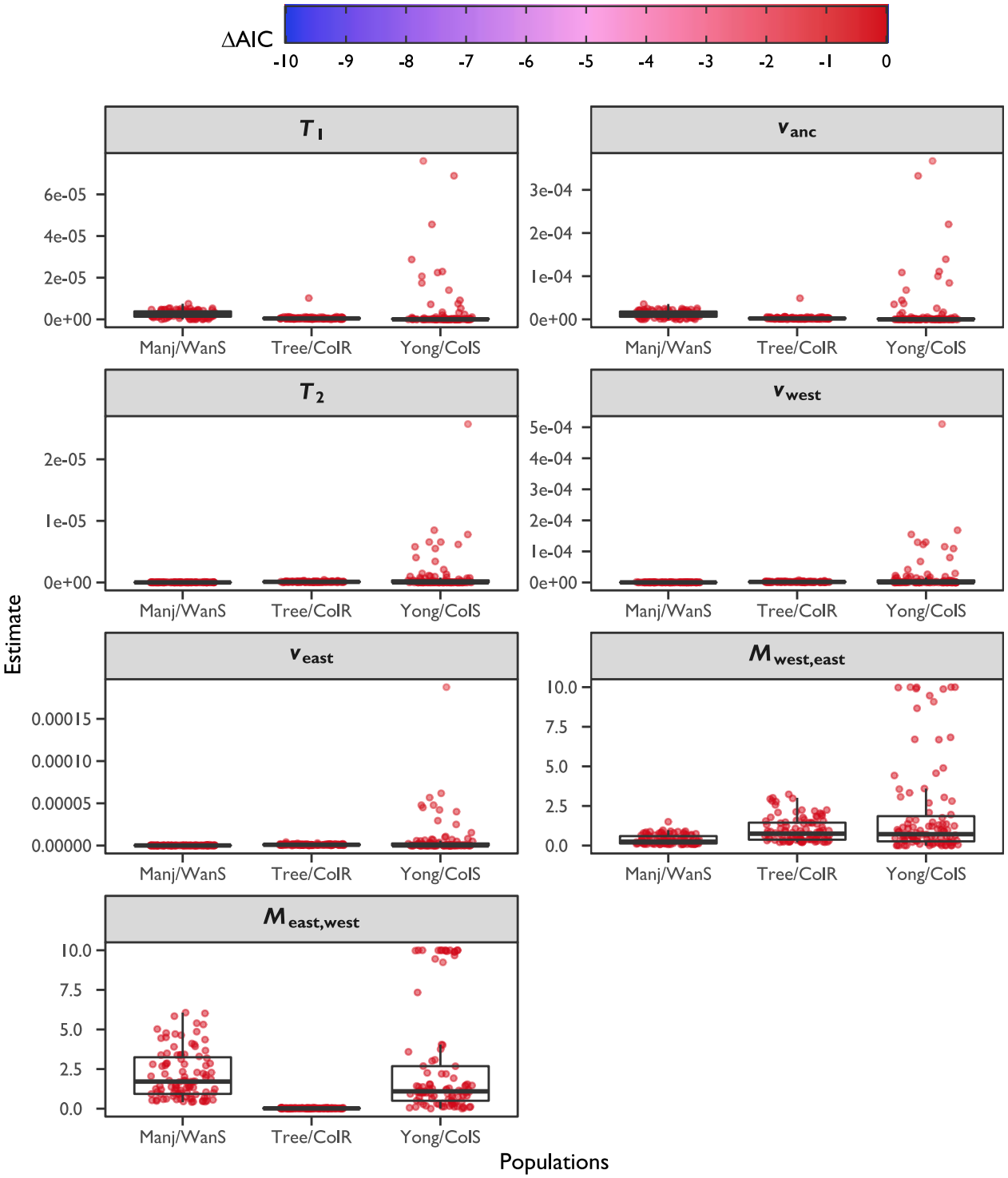

Figure S8. Parameter estimates across comparable replicate simulations (ΔAIC > –10) for the asymmetric migration scenario. Each panel represents a parameter. Population pairs are on the *x*-axis and parameter estimates are on the *y*-axis. Points represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend). Boxplots summarise the distribution of estimates. Population abbreviations: 'Manj', Manjimup; 'WanS', Wantirna South; 'Tree', Treeton; 'ColR', Colbinabbin Res; 'Yong', Yongarillup; 'ColS', Colbinabbin Sus.

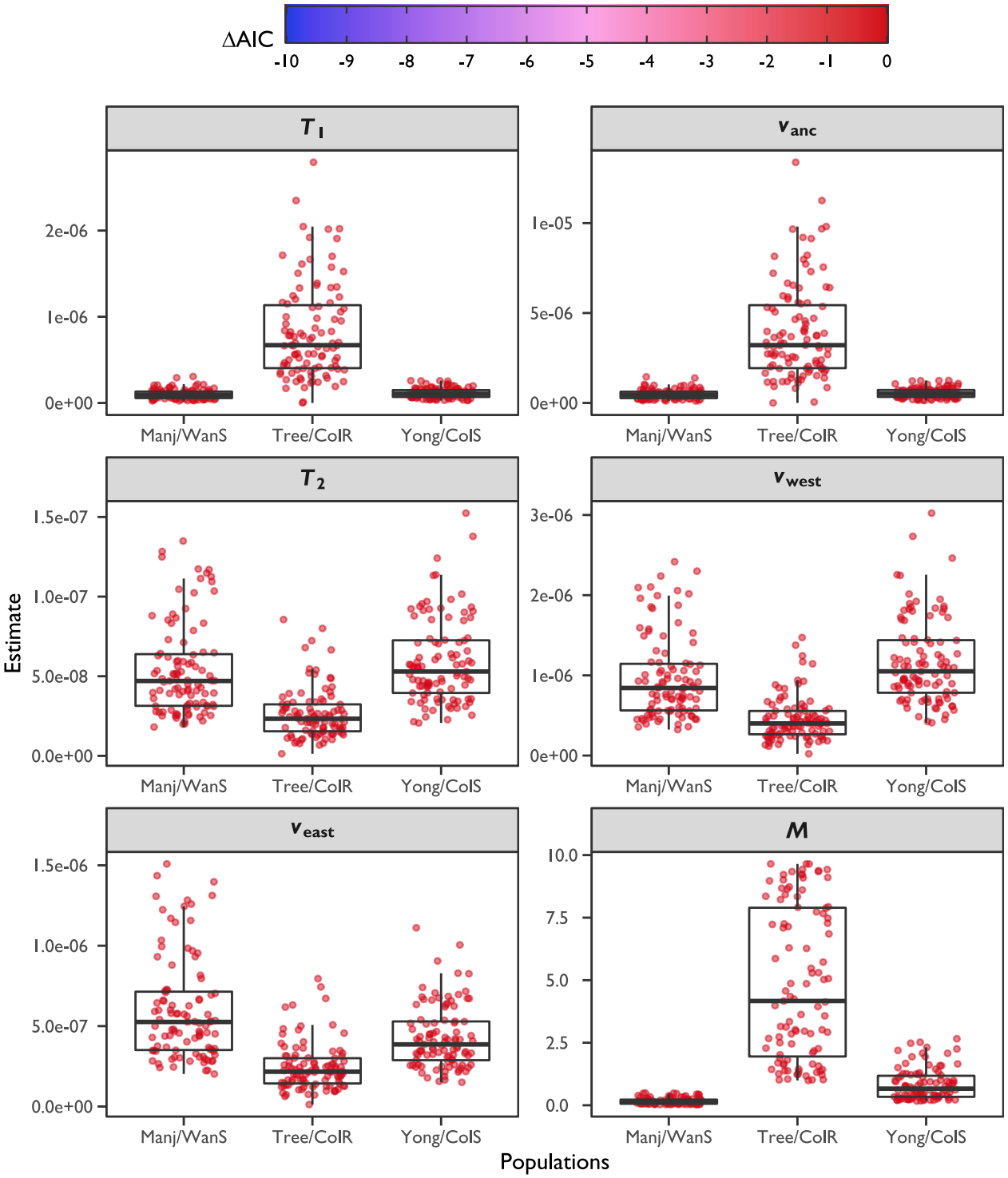

Figure S9. Parameter estimates across comparable replicate simulations (ΔAIC > –10) for the symmetric migration scenario. Each panel represents a parameter. Population pairs are on the *x*-axis and parameter estimates are on the *y*-axis. Points represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend). Boxplots summarise the distribution of estimates. Population abbreviations: 'Manj', Manjimup; 'WanS', Wantirna South; 'Tree', Treeton; 'ColR', Colbinabbin Res; 'Yong', Yongarillup; 'ColS', Colbinabbin Sus.

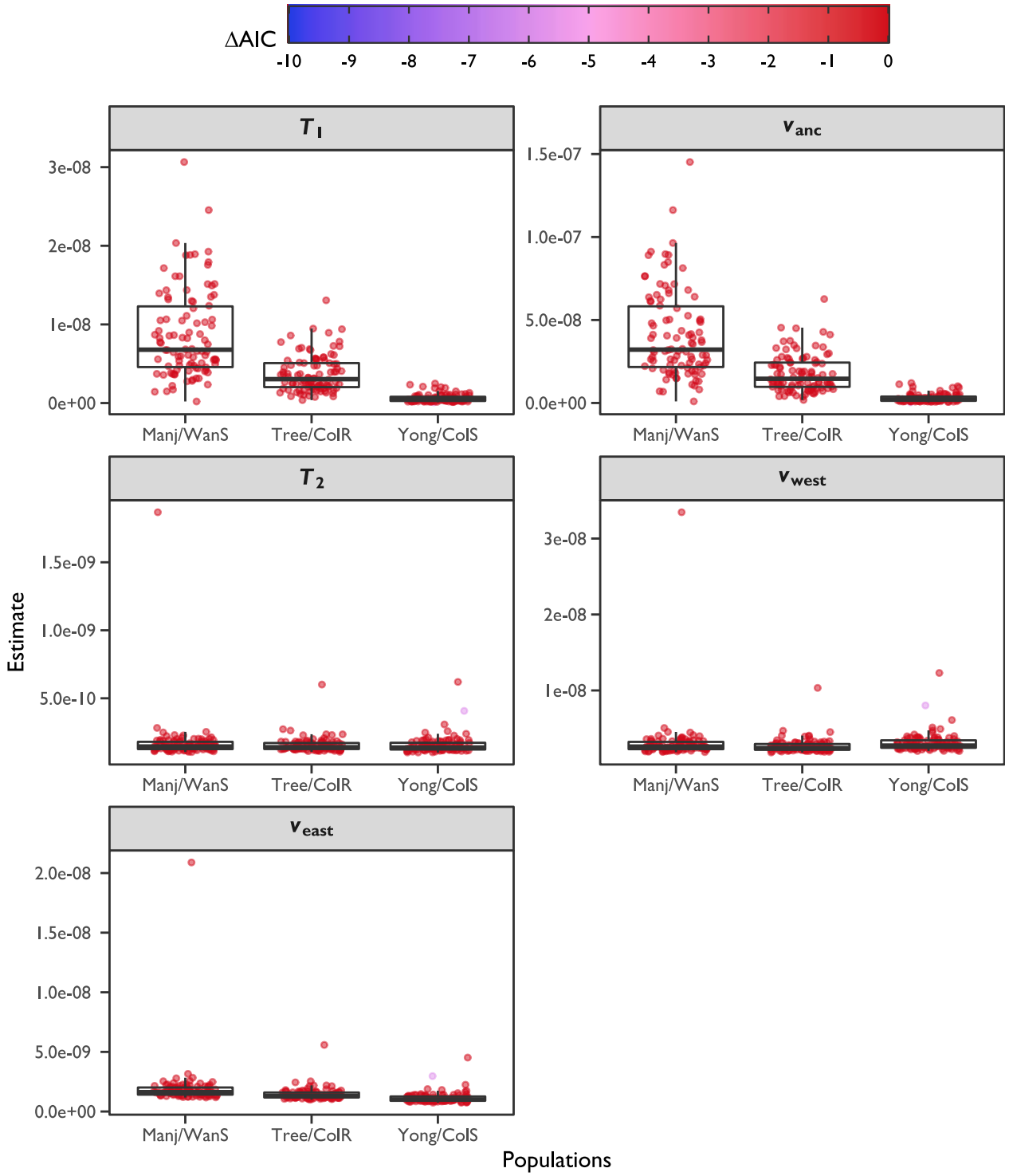

Figure S10. Parameter estimates across comparable replicate simulations (ΔAIC > –10) for the isolation scenario. Each panel represents a parameter. Population pairs are on the *x*-axis and parameter estimates are on the *y*-axis. Points represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend). Boxplots summarise the distribution of estimates. Population abbreviations: 'Manj', Manjimup; 'WanS', Wantirna South; 'Tree', Treeton; 'ColR', Colbinabbin Res; 'Yong', Yongarillup; 'ColS', Colbinabbin Sus.

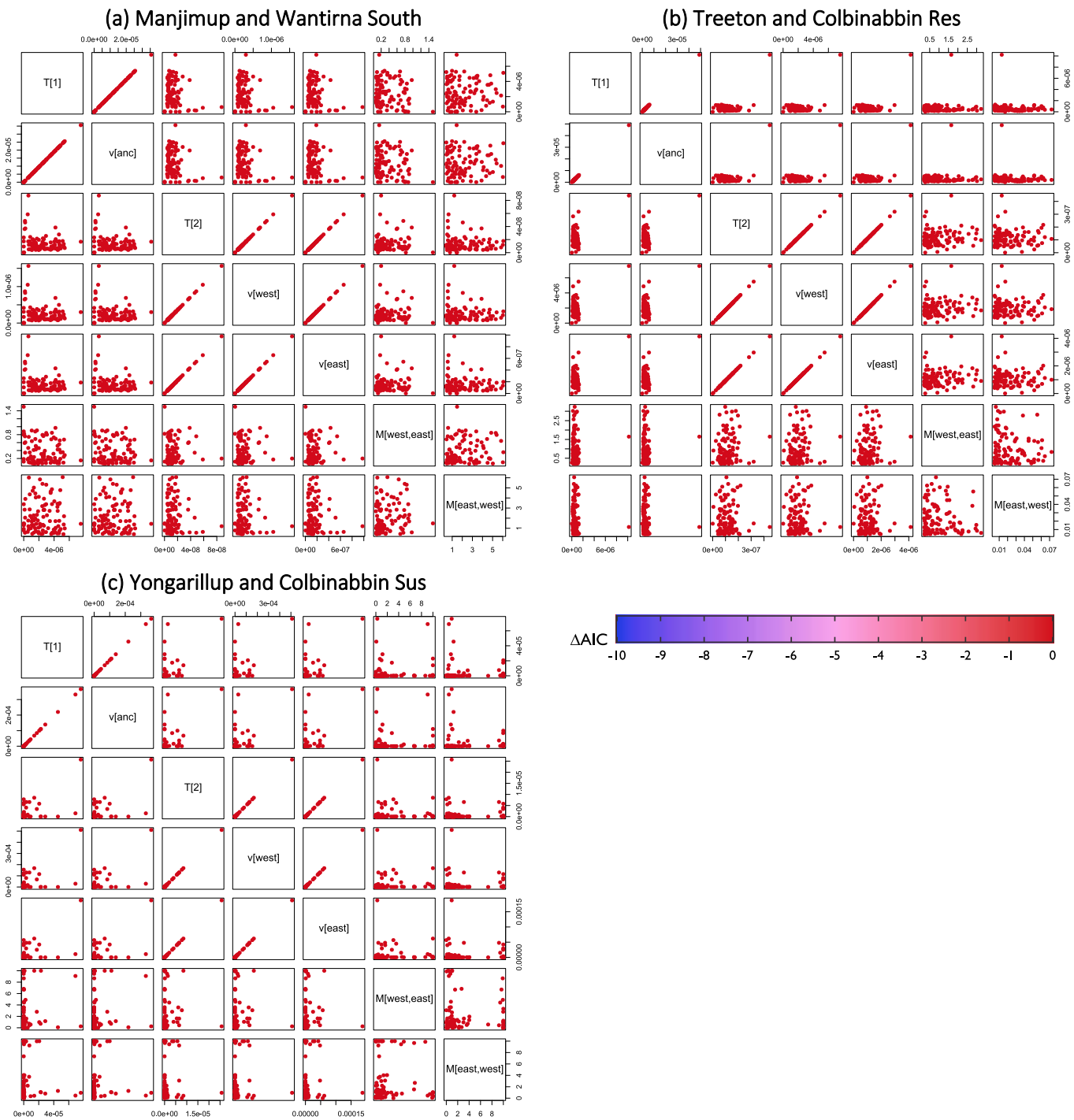

Figure S11. Correlation among parameter estimates for comparable replicate simulations (ΔAIC > –10) modelled under the asymmetric migration scenario, for each population pair. Points represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend).

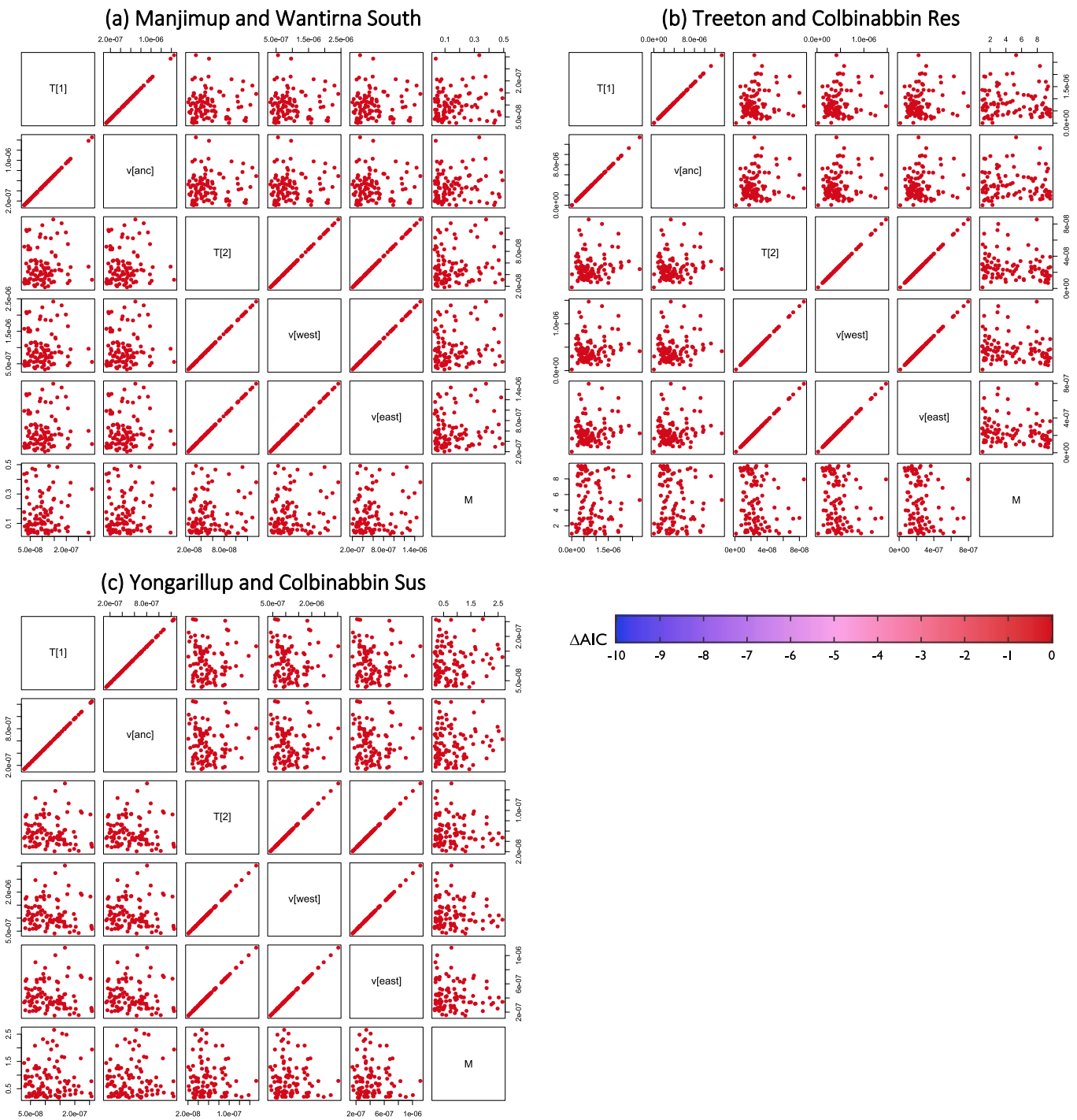

Figure S12. Correlation among parameter estimates for comparable replicate simulations (ΔAIC > –10) modelled under the symmetric migration scenario, for each population pair. Circles represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend).

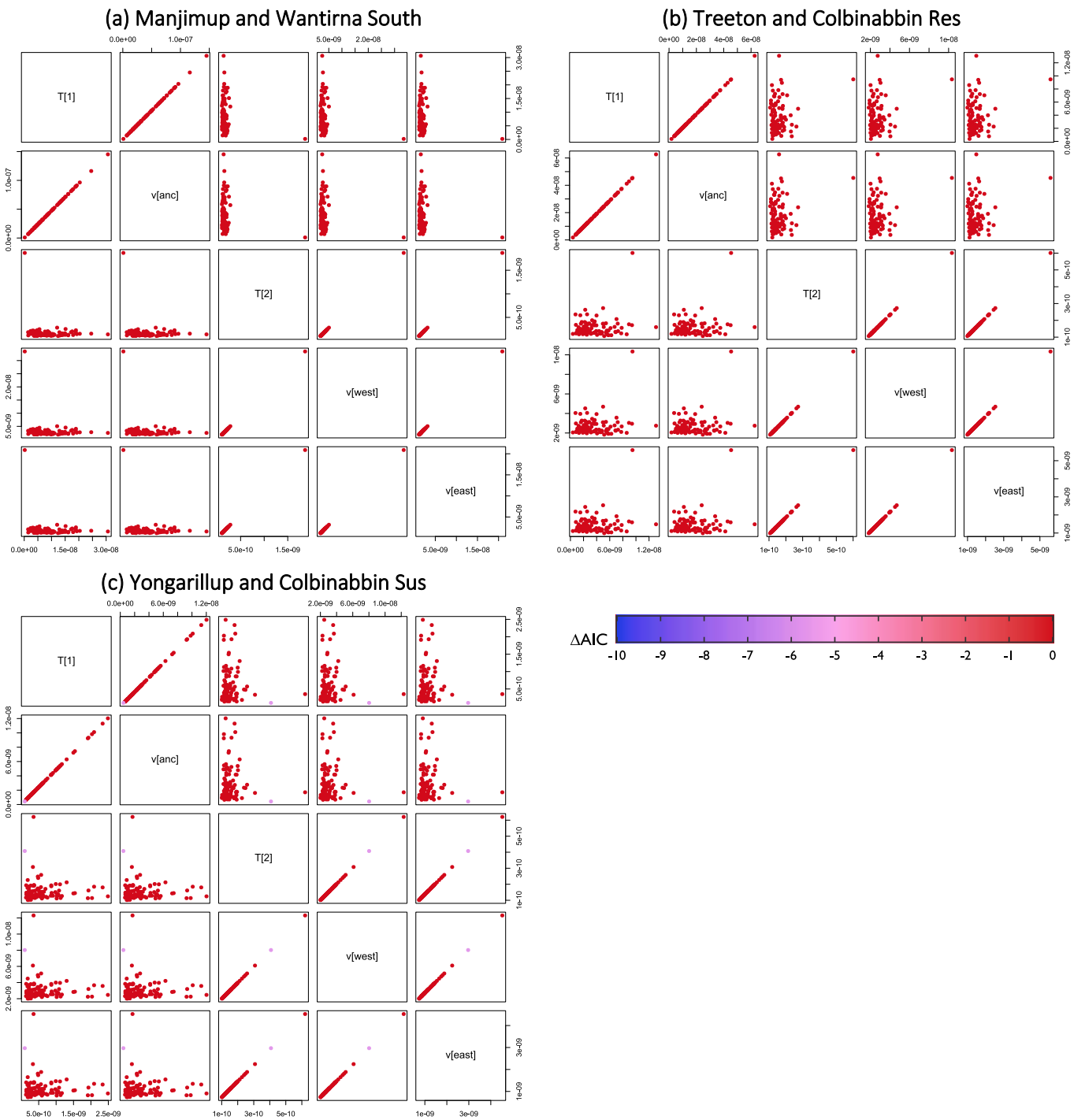

Figure S13. Correlation among parameter estimates for comparable replicate simulations (ΔAIC > –10) modelled under the isolation scenario, for each population pair. Points represent individual replicate simulations, coloured based on their ΔAIC to the best replicate simulation for their population pair (see legend).

Table S1. Samples used in population genomics analyses using whole-genome pool-seq of adult mites.

Sample ID	GenBank BioSample	Population (Library ID; Replicate ID)	Latitude	Latitude	Longitude	Pooled adult mites
RLEM2_2_A	SAMN30107459	Colbinabbin Sus (Lib=2; Rep=A)	2021	−36.5901	144.7994	50
RLEM2_2_B	SAMN30107460	Colbinabbin Sus (Lib=2; Rep=B)	2021	−36.5901	144.7994	50
RLEM4_1_A	SAMN30107463	Colbinabbin Res (Lib=4; Rep=A)	2021	−36.5642	144.8508	50
RLEM4_1_B	SAMN30107464	Colbinabbin Res (Lib=4; Rep=B)	2021	−36.5642	144.8508	50
RLEM18_X_A	SAMN30107483	Tambellup (Lib=18; Rep=A)	2019	−34.0456	117.5414	20
RLEM18_X_B	SAMN30107484	Tambellup (Lib=18; Rep=B)	2019	−34.0456	117.5414	20
RLEM19_1_A	SAMN30107485	Tambellup (Lib=19; Rep=A)	2019	−34.0456	117.5414	25
RLEM19_1_B	SAMN30107486	Tambellup (Lib=19; Rep=B)	2019	−34.0456	117.5414	25
RLEM20_1_A	SAMN30107487	Treeton (Lib=20; Rep=A)	2019	−33.851	115.2344	24
RLEM20_1_B	SAMN30107488	Treeton (Lib=20; Rep=B)	2019	−33.851	115.2344	24
RLEM21_1_A	SAMN30107489	Treeton (Lib=21; Rep=A)	2019	−33.851	115.2344	18
RLEM21_1_B	SAMN30107490	Treeton (Lib=21; Rep=B)	2019	−33.851	115.2344	18
RLEM26_2_A	SAMN30107491	Yongarillup (Lib=26; Rep=A)	2019	−33.7567	115.3869	14
RLEM26_2_B	SAMN30107492	Yongarillup (Lib=26; Rep=B)	2019	−33.7567	115.3869	14
RLEM27_2_A	SAMN30107493	Yongarillup (Lib=27; Rep=A)	2019	−33.7567	115.3869	29
RLEM27_2_B	SAMN30107494	Yongarillup (Lib=27; Rep=B)	2019	−33.7567	115.3869	29
RLEM28_2_A	SAMN30107495	Yongarillup (Lib=28; Rep=A)	2019	−33.7567	115.3869	25
RLEM28_2_B	SAMN30107496	Yongarillup (Lib=28; Rep=B)	2019	−33.7567	115.3869	25
RLEM29_2_A	SAMN30107497	Yongarillup (Lib=29; Rep=A)	2019	−33.7567	115.3869	14
RLEM29_2_B	SAMN30107498	Yongarillup (Lib=29; Rep=B)	2019	−33.7567	115.3869	14
RLEM34_1_A	SAMN30107503	Manjimup (Lib=34; Rep=A)	2019	−34.3101	116.1718	16
RLEM34_1_B	SAMN30107504	Manjimup (Lib=34; Rep=B)	2019	−34.3101	116.1718	16
RLEM35_X_A	SAMN30107505	Manjimup (Lib=35; Rep=A)	2019	−34.3101	116.1718	16
RLEM35_X_B	SAMN30107506	Manjimup (Lib=35; Rep=B)	2019	−34.3101	116.1718	16
RLEM36_1_A	SAMN30107507	Manjimup (Lib=36; Rep=A)	2019	−34.3101	116.1718	12

RLEM36_1_B	SAMN30107508	Manjimup (Lib=36; Rep=B)	2019	−34.3101	116.1718	12
RLEM37_X_A	SAMN30107509	Manjimup (Lib=37; Rep=A)	2019	−34.3101	116.1718	12
RLEM37_X_B	SAMN30107510	Manjimup (Lib=37; Rep=B)	2019	−34.3101	116.1718	12