# Genetic diversity in invasive populations [of Argentine stem weevil] allows adaptation to biocontrol

Thomas W.R. Harrop1, Marissa F. Le Lec1, Ruy Juaregui2, Shannon Taylor1, Sarah Inwood1, John Skelly1, Siva Ganesh (**sp?**)2, Rachael Ashby3, Jeanne Jacobs3, Stephen Goldson4, Peter K. Dearden1

**Goldson lab dissection ppl? Stephen, please add.**

1 University of Otago & GA  
2 AgResearch Palmerston North  
3 AgResearch Lincoln?  
4 BPRC

## Abstract

* context and need for the work
  + Modified, agricultural landscapes are susceptible to damage by insect pests.
  + Biological control of pests is typically successful once the control agent is established, but this depends on its ability to adapt to evolution in the host.
  + Theoretical studies have shown that unequal genetic variation between the host and the control agent will lead to rapid evolution of resistance in the host, but cases of this have not been documented in the field.
* approach and methods used
  + To address this, we measured the genetic variation in New Zealand populations of the pasture pest, Argentine stem weevil (*Listronotus bonariensis*), which is controlled with declining effectiveness by a parasitoid wasp, *Microctonus hyperodae*.
  + We constructed a draft reference genome, then collected samples from a geographical survey of 10 sites around New Zealand and genotyped them using a modified genotyping-by-sequencing approach.
* main results (2-3 points)
  + New Zealand populations of Argentine stem weevil have high levels of heterozygosity and low levels of inbreeding, consistent with a large effective population size and frequent gene flow.
  + This implies that Argentine stem weevils were able to evolve more rapidly than its biocontrol agent, which reproduces asexually.
  + These findings show that monitoring genetic diversity in biocontrol agents and their targets is critical for long-term success of biological control.

### Synthesis and applications

* wider implications and relevance to management or policy

### Keywords

Naughty weevils, Biocontrol, Invasive species, Molecular evolution, ???

## Introduction

Biocontrol paragraph (not written yet):

* use of biocontrol
* problems with establishment
* likelihood / incidence of resistance after establishment
* Once established, failure of biological control is rare/unprecedented (?), because **xyz** (**ref**).

New Zealand pastures are highly modified landscapes that suffer from severe pest impacts [1]. The susceptibility of pasture to pests may be due to low plant and animal diversity, resulting in low biotic resistance to invasive species [2]. The Argentine stem weevil (ASW), *Listronotus bonariensis* Kuschel (Coleoptera: Curculionidae), is a particularly destructive, invasive pest of pasture, reaching densities of 700 adults per m2 and causing economic impacts up to NZ$200M per annum [1,3,4]. Conventional, chemical control of ASW is not effective because **xyz** (**ref**). To complement endophyte-based plant resisance [5,6], the solitary wasp *Microctonus hyperodae* Loan (Hymenoptera: Braconidae) was released for biological control of *L. bonariensis* in 1992. Within three years of its release, parasitism of ASW by *M. hyperodae* reached 90% [7], reducing or eliminating damage by ASW [7–9].

Although ASW was initially well managed by this system, biological control of ASW by *M. hyperodae* failed after about 14 generations [10–12]. This failure may be the result of adaptation in weevil populations resulting from selection pressure by the parasitoid [11,12]. Because ASW reproduces sexually, ASW populations may have greater capacity to evolve than populations of *M. hyperodae*, which reproduces parthenogenetically. Empirical modelling of the ASW–*M. hyperodae* interaction indicated that resistance is inevitable when hosts have more genetic variation than their predator [13], *e.g.* because of genetic recombination. Although a theoretical pathway for resistance has been established, examples of evolution of resistance to classical biological control have not been reported (**are we SURE sure?**).

Measuring genetic variation in populations of the host and the parasitoid is required to explain this case of evolution of resistance to biocontrol. We address this by genotyping-by-sequencing of a geographical survey of Argentine stem weevil populations in New Zealand. Our experiments reveal a high heterozygosity and a high proportion of unstructured variation across 12 populations from the North and Sourth islands of New Zealand, consistent with large effective population size and gene flow between populations. Genetic variation along a latitudinal cline is associated with signatures of selection in regions of the genome, indicating a level of local adaptation within populations, but we found no evidence at this resolution of genetic adaptation in parasitised weevils compared to parasitoid-free weevils. Our results show that the amount of genetic variation in NZ populations of ASW is far higher than detected by traditional molecular markers [14,15], suggesting that ASW populations evolved resistance via weak selection acting on variants of minor effect that existed before the introduction of *M. hyperodae*.

## Materials and methods

### Weevil sampling

**Stephen, I need these details from your team please**:

* weevil collection details for geographic survey
* collection and processing/dissection details for parasitised *vs*. unparasitised expt

The map in Figure 1 was plotted with the ggmap package for ggplot2 [16].

### Genome assembly

To produce the short read dataset, an Illumina TruSeq PCR-free 350bp insert library was generated from DNA extracted from a single, male Argentine stem weevil collected from endophyte-free hybrid ryegrass (*Lolium perenne* × *Lolium multiflorum*) at Lincoln, New Zealand. Library preparation and sequencing were performed by Macrogen Inc. (Seoul, Republic of Korea). A total of 158 GB of 100 b and 150 b paired-end reads were generated from the TruSeq PCR-free library. After removing common sequencing contaminants and trimming adaptor sequences using BBTools [17], the short-read-only genome was assembled with meraculous [18–20]. Reproducible code for assembling the short-read dataset and assessing the assemblies is hosted at [github.com/tomharrop/asw-nopcr](https://github.com/tomharrop/asw-nopcr).

To produce long reads from a single individual, we produced high molecular weight DNA from a single, male ASW collected from Ruakura, New Zealand, using a modified QIAGEN Genomic-tip 20/G extraction protocol [21]. We amplified the DNA using Φ29 multiple displacement amplification (QIAGEN REPLI-g Midi Kit) and debranched the amplified DNA using T7 Endonuclease I (New England Biolabs) according to the Oxford Nanopore Technologies Premium whole genome amplification protocol version WGA\_kit9\_v1. Debranching reduced the raw read *N*50 length to 9.0 KB. Amplified DNA was sequenced on 6 R9.4.1 flowcells using a MinION Mk1B sequencer (Oxford Nanopore Technologies). We also extracted high molecular weight DNA from three pools, each of 20 unsexed individuals collected from Ruakura, New Zealand. We sequenced this pooled DNA on 5 R9.4.1 flowcells, following the Genomic DNA by Ligation protocol (SQK-LSK109; Oxford Nanopore Technologies). We removed adaptor sequences from the long reads with Porechop 0.2.4 ([github.com/rrwick/Porechop](https://github.com/rrwick/Porechop)) and assembled with Flye 2.6 [22]. Reproducible code for assembling and assessing the long-read ASW genomes is hosted at [github.com/TomHarrop/asw-flye-withpool](https://github.com/TomHarrop/asw-flye-withpool).

All genome assemblies were assessed by size and contiguity statistics and BUSCO analysis [23]. Redundant contigs were removed from the combined, long read assembly with Purge Haplotigs 0b9afdf [24] using a low, mid and high cutoff of 60, 120 and 190, respectively. We attempted to use RepeatModeler 2.0.1 [25] and RepeatMasker 4.1.0 [26] from the Dfam TE Tools Container v1.1 ([github.com/Dfam-consortium/TETools](https://github.com/Dfam-consortium/TETools)) to estimate the repeat content of the long read genomes, but >500M high-scoring Segment Pairs (HSPs) were identified and RepeatModeler did not finish after running for 4 weeks with 144 GB of physical RAM.

### Reduced-representation genome sequencing, processing and analysis

**Jeanne, I need these details from your team please**:

* details on DNA extraction, GBS pipeline and sequencing

All the code we used to process the raw reads, assemble loci and run downstream analyses is hosted at [github.com/TomHarrop/stacks-asw](https://github.com/TomHarrop/stacks-asw), including the parameters and software containers for each step.

Briefly, we used a strict processing pipeline to prepare the raw GBS reads for locus assembly. Samples were demultiplexed with zero allowed barcode mismatches to 91–93 b reads, depending on barcode length. Reads were trimmed by searching for adaptors with a minimum match of 11 b. Reads shorter than 80 b after trimming were discarded. All remaining reads were truncated to 80 b to account for unmatched adaptor sequence < 11 b that may have been present at the end of reads. To remove overamplified samples, we calculated the GC content for each library and discarded samples with median read GC > 45%. We assembled loci against our draft genome using gstacks 2.53 [27].

For analysis, we used bcftools [ref?] to remove sites with more than 2 alleles, minor allele frequency < 0.05, or missing genotypes in more than 20% of individuals. After filtering loci, we also removed individuals that were missing genotypes at more than 20% of loci. We ran the Stacks 2.53 populations module [27] to calculate inbreeding (*F*) and heterozygosity statistics. We used plink 1.9 [28] to prune sites in linkage disequilibrium for principal components analysis and discriminant analysis of principal components with the adegenet 2.1.2 package for R [29,30]. We used PGDSpider 2.1.1.5 [31] to convert the un-pruned dataset for detection of loci under selection with BayeScan 2.1 [32]. After statistically phasing SNPs from the un-pruned dataset with SHAPEIT v2 r904 [33], we analysed cross-population extended haplotype homozygosity with the R package rehh 3.1.0 [34]. For demographic analysis, we converted the pruned dataset to site frequency spectra using easysfs commit c2b26c5 from [github.com/isaacovercast/easySFS](https://github.com/isaacovercast/easySFS), and tested demographic models using **fastsimcoal2 2.6 (tbc)** [35].

### Reproducibility and data availability

Raw sequence data for the ASW genome are hosted at the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) under accession **TBA**. We used snakemake [36] to arrange analysis steps into workflows and monitor dependencies, and Singularity [37] to capture the computing environment. Using the code repositories listed in each methods section, the final results can be reproduced from the raw data with a single command using snakemake and Singularity. The source for this manuscript is hosted at [github.com/TomHarrop/asw-gbs-genome-paper](https://github.com/TomHarrop/asw-gbs-genome-paper).

## Results

### The Argentine stem weevil genome is repetitive and polymorphic

To construct a reference for genotyping populations of Argentine stem weevils, we produced a draft assembly of the ASW genome. We initially attempted assembly from a single individual using PCR-free, short read sequencing. This resulted in a fragmented assembly with low BUSCO scores (Table 1). *k*-mer analysis on the raw short reads suggested 2.1 polymorphisms per 100 bp and a genomic repeat content of at least 28% in the individual we sequenced (**Supporting Information**). We then attempted to produce a long-read genome assembly using whole-genome amplification (WGA) of high molecular weight (HMW) DNA from a single individual, followed by sequencing on the Oxford Nanopore Technologies (ONT) MinION sequencer. We produced 29.8 GB of quality-filtered reads with an *N*50 length of 9.0 KB. Assembling the single individual, long read genome resulted in improved contiguity and BUSCO scores compared to the short-read assembly (Table 1). Consistent with the raw short read data, we detected an **extreme level (how much? RM isn’t going to finish)** of repeats in the single individual, long read genome (Table 1). To improve assembly across long repeats, we produced a second ONT dataset with longer reads from HMW DNA from two pools of 20 individuals each, without amplification. Sequencing these samples on the MinION sequencer produced a total of 12.0 GB of quality-filtered reads with an *N*50 length of 19.5 KB. Assembling the longer reads generated from the pooled sample alone resulted in a more contiguous genome, but with lower BUSCO scores (Table 1). We constructed a combined, long-read genome using the pooled, long-read dataset for contig construction, and the single-individual, long-read dataset for assembly polishing. This improved the BUSCO scores, but produced a large number of redundant contigs (Table 1), presumably because of the high rate of heterozygosity in the pooled, long-read dataset. Finally, we used the PCR-free, short read sequencing data from a single individual with the Purge Haplotigs pipeline to remove redundant contigs from the combined long read assembly [24]. This resulted in a final draft assembly of 1.1 GB with an *N*50 length of 122.3 kb and a BUSCO completeness of 83.9%.

**Table 1**. Assembly statistics for the final draft genome and intermediate assemblies. n.d.: not determined. x: RepeatModeler has been running for 6 weeks now, it’s never going to finish.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Short read | Single individual, long read | Pooled, long read | Combined, long read | Final draft |
| Assembly length (Gb) | 1.3 | 1.2 | 1.2 | 1.7 | 1.1 |
| *N*50 | 53046 | 4523 | 2958 | 5281 | 2681 |
| *N*50 length (kb) | 7.1 | 74.4 | 112.6 | 86.4 | 122.3 |
| Complete, single-copy BUSCOs (%) | 32.7 | 72.2 | 71.0 | 69.2 | 78.8 |
| Complete, multiple-copy BUSCOs (%) | 17.2 | 7.5 | 5.9 | 17.4 | 5.1 |
| Repeat content (%) | n.d. | x | x | x | ~67.8 |

### Genetic variation is associated with geography in NZ populations of Argentine stem weevil

To measure genetic variation in invasive New Zealand populations of ASW, we collected individuals from 10 sites across the North and South Islands of New Zealand (Figure 1A). We genotyped individuals with a modified genotyping-by-sequencing (GBS) protocol [38]. After strict trimming and filtering of the raw GBS data, we mapped reads from each individual against our draft genome and used gstacks to assemble loci [27]. For analysis, we removed loci with more than two alleles, minor allele frequency less than 0.05, or missing genotypes in more than 20% of individuals. We also removed individuals missing genotypes at more than 20% of loci. The complete dataset comprised 7–15 individuals per location (total 116), genotyped at 52,051 biallelic SNPs. The mean observed heterozygosity ranged from 0.18–0.21 across populations (Figure 1B), and pairwise *F*ST values between populations ranged from 0.024–0.051 (Figure 1C). For principal components analysis (PCA), we pruned the dataset to 18,715 biallelic SNPs that were not in linkage disequilibrium, using an r2 threshold of 0.1. PCA of genotypes at these sites revealed overlapping populations of ASW, with 9.2% of total variance explained by the first two components (Figure 1D). These results suggest that NZ populations of ASW are highly heterozygous, but the variation is not highly structured between populations, consistent with a large effective population size and high gene flow between populations. We used discriminant analysis of principal components (DAPC) on the same set of pruned SNPs to find genetic variability associated with differences between populations [29]. The major linear discriminant, which explains 96.7% of between-population variation, separates populations from North and South of the Alpine divide (Figure 1E), although admixture was evident in all populations except Lincoln (Figure 1F). This indicates a degree of genetic isolation between populations from North and South of the Alpine divide.

![Figure 1. Caption next page.](data:application/pdf;base64,)

**Figure 1.** Caption next page.

**Figure 1.** Genetic diversity in NZ populations of Argentine stem weevil. **A** Weevil sampling locations. We collected Argentine stem weevils from 4 locations in the North Island and 6 locations in the South Island of New Zealand. Greymouth is in the South island, but North of the Alpine divide. The number of weevils genotyped from each location is shown on the map. **B** Mean observed heterozygosity for each population. **C** Pairwise *F*ST values between populations. **D** Principal components analysis (PCA) and **E** discriminant analysis of principle components (DAPC) of 116 individuals genotyped at 18,715 biallelic sites. **D** The populations overlap on the first two principal components (PC1 and PC2), but weevils sampled from higher latitudes have lower scores on PC1. PC1 and PC2 together explain 9.2% of variance in the dataset, indicating a high level of unstructured genetic variation in weevil populations. **E** Linear discriminant 1 (LD1) explains 96.7% of between-group variability. Individuals from North of the Alpine divide have negative coordinates on LD1, whilst individuals from South of the Alpine divide have positive coordinates. LD2 separates Lincoln individuals from other individuals. **F** Posterior probability of group assignment for each individual. All populations contain individuals with high posterior probabilities of membership to other populations, consistent with admixture (**?**). We did not detect admixture between populations from North and South of the Alpine divide. Individuals sampled from Lincoln had the lowest posterior probabilities of membership to other populations.

### Genetic variation is not associated with parasitism by *M. hyperodae*

To detect large-effect variants associated with susceptibility to parasitism by *M. hyperodae*, we genotyped weevils that had also been tested for the presence of a parasitoid larva. We used a total of 179 individuals, collected from Lincoln, New Zealand, and Ruakura, New Zealand, because of the decline in parasitism rate recorded at these locations [12]. The weevils were examined for a parasitoid larva and genotyped at the same loci used for the geographical survey. After filtering and pruning sites in linkage disequilibrium, we used 19,482 SNPs for PCA and DAPC. We did not detect any genetic differentiation associated with the presence of a parasitoid, either within populations or between populations, or any evidence of skewed allele frequencies in these groups using BayeScan (lowest *Q*-value 0.97).

@ Marissa and Peter - there are some contigs that come up using XPEHH for these groups. There’s a contig (contig\_11238) full of transcription factors with a single locus that has a high positive XPEHH score, but this would suggest selection in the *parasitised* group. There is one contig (contig\_5955) with a large negative XPEHH (selection in non-parasitised group) but the genes are not interesting. I think the XPEHH results might be spurious because the low number of markers per contig lead to dodgy phasing, so I’m thinking of removing XPEHH from the following section. I’ve left them in this version to get your thoughts.

### Genetic differentiation between ASW populations North and South of the Alpine divide

Although we did not detect variation associated with presence of a parasitoid, parasitism rate varies across sites in NZ [12]. To investigate the genetic differentiation between regions, we grouped individuals that were collected from North and South of the Alpine divide (Figure 1). The two groups had a mean *F*ST of 0.013. We detected 47 SNPs with skewed allele frequencies across 24 contigs in the draft genome with BayeScan (Figure 2). The contigs containing these SNPs had a total of 3–36 SNPs, and all 47 of the detected SNPs had positive α values, suggesting diversifying selection (Table 2). Using an orthogonal method, 32 SNPs across 5 contigs had outlying cross-population extended haplotype homozygosity (XPEHH) scores [39,40]. Both methods identified putative SNPs under selection an overlapping region on contig\_40523. These sites had high α values and positive XPEHH scores, suggesting diversifying selection in the North group. We identified four genes on contig\_40523. None had characterized functions in insects.

![Figure 2. A Regions of the draft ASW genome that have altered allele frequencies between populations from North and South of the Alpine divide. 47 SNPs on 24 contigs have altered allele frequencies, using the arbitrary Q-value cutoff of 0.01. B Models of population demographics and results (in progress).](data:application/pdf;base64,)

**Figure 2**. **A** Regions of the draft ASW genome that have altered allele frequencies between populations from North and South of the Alpine divide. 47 SNPs on 24 contigs have altered allele frequencies, using the arbitrary *Q*-value cutoff of 0.01. **B** Models of population demographics and results (in progress).

**Table 2**. Number of SNPs under selection using BayeScan [32] (*Q* < 0.01) or cross-population extended haplotype homozygosity (XPEHH) analysis [39,40] (-log10*p* > 4). α is BayeScan’s locus-specific component of *F*ST coefficient [32]. Positive values suggest diversifying selection. Positive XPEHH scores suggest selection in the North group, and negative scores suggest selection in the South group.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Contig | Total SNPs | BayeScan SNPs | BayeScan region | α | XPEHH SNPs | XPEHH region | XPEHH |
| contig\_40523 | 26 | 5 | 103,989 – 111,755 | 1.93 – 2.14 | 2 | 111,724 – 111,783 | 5.34 – 5.83 |
| contig\_11164 | 23 | 4 | 60,487 – 179,797 | 1.84 – 2.05 | 0 |  |  |
| scaffold\_43207 | 11 | 4 | 102,783 – 102,811 | 1.70 – 2.14 | 0 |  |  |
| contig\_2677 | 22 | 3 | 109,260 – 110,069 | 1.79 – 1.99 | 0 |  |  |
| contig\_18336 | 20 | 3 | 233,287 – 342,618 | 1.90 – 1.99 | 0 |  |  |
| contig\_12006 | 10 | 3 | 46,975 – 47,031 | 2.09 – 2.15 | 0 |  |  |
| contig\_13287 | 10 | 3 | 58,727 – 58,738 | 1.68 – 1.98 | 0 |  |  |
| contig\_39072 | 10 | 3 | 55,060 – 55,126 | 2.06 – 2.14 | 0 |  |  |
| contig\_37676 | 17 | 2 | 47,912 – 47,954 | 2.01 – 2.03 | 0 |  |  |
| contig\_4080 | 13 | 2 | 26,024 – 26,027 | 1.78 – 1.79 | 0 |  |  |
| contig\_23638 | 6 | 2 | 118,602 – 118,616 | 2.08 – 2.15 | 0 |  |  |
| contig\_8456 | 36 | 1 | 88,819 | 1.92 | 0 |  |  |
| contig\_1196 | 24 | 1 | 393,177 | 1.68 | 0 |  |  |
| contig\_20252 | 20 | 1 | 34,278 | 1.87 | 0 |  |  |
| contig\_27115 | 20 | 1 | 111,838 | 1.94 | 0 |  |  |
| contig\_3057 | 13 | 1 | 272,534 | 1.91 | 0 |  |  |
| contig\_23312 | 12 | 1 | 171,714 | 1.95 | 0 |  |  |
| contig\_202 | 8 | 1 | 80,251 | 2.03 | 0 |  |  |
| contig\_14933 | 7 | 1 | 80,478 | 1.98 | 0 |  |  |
| contig\_19450 | 6 | 1 | 37,951 | 2.39 | 0 |  |  |
| contig\_205 | 6 | 1 | 28,713 | 2.13 | 0 |  |  |
| contig\_21253 | 6 | 1 | 43,048 | 1.76 | 0 |  |  |
| contig\_28985 | 6 | 1 | 22,034 | 2.01 | 0 |  |  |
| contig\_12091 | 3 | 1 | 43,431 | 2.02 | 0 |  |  |
| contig\_1525 | 45 | 0 |  |  | 18 | 30,186 – 47,858 | -9.56 – -5.17 |
| contig\_5179 | 41 | 0 |  |  | 9 | 46,573 – 48,037 | -7.02 – -5.48 |
| contig\_18 | 52 | 0 |  |  | 2 | 488,821 – 488,823 | -6.85 – -5.39 |
| contig\_71 | 22 | 0 |  |  | 1 | 247,010 | 7.34 |

### Repeated/large/separate incursions of ASW into New Zealand

This is in progress. We’re testing 3 main models:

* model 1: single introduction then bottleneck, spread, diversification, and gene flow.
* model 2: **separate introductions** to North and South, with bottlenecks and gene flow between populations.
* model 3: separate introductions, then bottlenecks, from **different source populations** to North and South, with bottlenecks and gene flow between populations.
* all of the above but without bottlenecks to simulate large or repeated incursions.

## Discussion

The purpose of this work was to investigate the relationship between genetic variation and resistance to biocontrol in New Zealand populations of Argentine stem weevil. Previous reports using randomly amplified polymorphic DNA (RAPD) markers and *cytochrome C oxidase subunit I* (*COI*) sequencing suggested a high degree of genetic similarity and identified a single *COI* haplotype in New Zealand populations [14,15]. In contrast, our results from a genome-wide genotype-by-sequencing (GBS) approach reveal a high level of genetic diversity within and between populations. We suggest that this standing variation provides an evolutionary advantage to ASW populations in comparison to their biocontrol agent, which we expect to be exacerbated by asexual reproduction in *M. hyperodae*. This indicates that genetic variation in both host and target need to be monitored with high-resolution genotyping to maintain success of biological control.

**Para on demographics goes here (or not)**.

Despite the increased resolution of GBS compared to traditional markers, we did not detect regions of the genome associated with parasitism by *M. hyperodae*. Resistance to biocontrol may not be genetic, or manual detection of the parasitoid may not be a strong enough phenotype to separate resistant and susceptible individuals. We suggest that resistance to *M. hyperodae* is encoded by multiple regions of small effect, selected from the high amount of standing variation we detected in ASW populations. A polygenic response on standing variation is consistent with the low level of selection by the biocontrol agent [41,42]. The number of markers yielded by legacy GBS approaches provide low power to detect polygenic responses resulting from weak selection on standing variation. Higher-resolution genome-wide association studies using genome-wide resequencing with more individuals and a stronger resistance phenotype may allow detection of variation associated with resistance of the weevils to biocontrol.

**Probably remove this para**. Our final genome is draft quality and we expect gaps in the assembly at larger repeat regions that were not sufficiently covered by long reads. Short-read genome assembly was not effective for this species because of the extreme repeat content. Our final draft assembly had a repeat content of **67.8%** (Table 1), with a maximum repeat size of 17.7 kb and a repeat *N*50 length of 485 bp. The non-repetitive regions had an *N*50 length of 1066 bp. The heterozygosity in weevil populations and lack of an inbred, laboratory strain made pooling individuals for sequencing undesirable. Our assembly strategy of contig construction with the longest reads, followed by assembly polishing with long reads from a single individual, and then redundant contig removal with PCR-free short reads from another single individual allowed us to improve the contiguity and completeness of the stem weevil genome (Table 1).

**Conclusion (let’s see what demographics say)**

## Authors’ contributions

## Acknowledgements

## Data availability

## References

1. Ferguson, C.M.; Barratt, B.I.P.; Bell, N.; Goldson, S.L.; Hardwick, S.; Jackson, M.; Jackson, T.A.; Phillips, C.B.; Popay, A.J.; Rennie, G. et al. Quantifying the economic cost of invertebrate pests to New Zealand’s pastoral industry. *New Zealand Journal of Agricultural Research* **2019**, *62*, 255–315. doi: [10.1080/00288233.2018.1478860](https://doi.org/10.1080/00288233.2018.1478860).

2. Goldson; Barker; Chapman; Popay, A.J.; Stewart; Caradus; Barratt, B.I.P. Severe insect pest impacts on New Zealand pasture. *Journal of Insect Science* **2020**, *in press*.

3. Barker, G.; Addison, P.J. Argentine stem weevil populations and damage in ryegrass swards of contrasting Acremonium infection. In Proceedings of the Proceedings of the 6th Australasian Conference on Grassland Invertebrate Ecology; 1993.

4. Prestidge, R.; Barker, G.; Pottinger, R. The economic cost of Argentine stem weevil in pastures in New Zealand. In Proceedings of the Proceedings of the 44th New Zealand Weed and Pest Control Conference; 1991; Vol. 44, pp. 165–170.

5. Johnson, L.J.; de Bonth, A.C.M.; Briggs, L.R.; Caradus, J.R.; Finch, S.C.; Fleetwood, D.J.; Fletcher, L.R.; Hume, D.E.; Johnson, R.D.; Popay, A.J. et al. The exploitation of epichloae endophytes for agricultural benefit. *Fungal Diversity* **2013**, *60*, 171–188. doi: [10.1007/s13225-013-0239-4](https://doi.org/10.1007/s13225-013-0239-4).

6. Kauppinen, M.; Saikkonen, K.; Helander, M.; Pirttilä, A.M.; Wäli, P.R. Epichloë grass endophytes in sustainable agriculture. *Nature Plants* **2016**, *2*, 1–7. doi: [10.1038/nplants.2015.224](https://doi.org/10.1038/nplants.2015.224).

7. Barker, G.M.; Addison, P.J. Early Impact of Endoparasitoid *Microctonus* *Hyperodae* (Hymenoptera: Braconidae) After Its Establishment in *Listronotus* *Bonariensis* (Coleoptera: Curculionidae) Populations of Northern New Zealand Pastures. *Journal of Economic Entomology* **2006**, *99*, 273–287. doi: [10.1093/jee/99.2.273](https://doi.org/10.1093/jee/99.2.273).

8. Goldson, S.L.; Barron, M.C.; Kean, J.M.; Koten, C. van Argentine stem weevil (*Listronotus* *Bonariensis*, Coleoptera: Curculionidae) population dynamics in Canterbury, New Zealand dryland pasture. *Bulletin of Entomological Research* **2011**, *101*, 295–303. doi: [10.1017/S0007485310000507](https://doi.org/10.1017/S0007485310000507).

9. Barker, G.M. Biology of the Introduced Biocontrol Agent *Microctonus* *Hyperodae* (Hymenoptera: Braconidae) and Its Host *Listronotus* *Bonariensis* (Coleoptera: Curculionidae) in Northern New Zealand. *Environmental Entomology* **2013**, *42*, 902–914. doi: [10.1603/EN11248](https://doi.org/10.1603/EN11248).

10. Popay, A.J.; McNeill, M.R.; Goldson, S.L.; Ferguson, C.M. The current status of Argentine stem weevil (*Listronotus* *Bonariensis*) as a pest in the North Island of New Zealand. *New Zealand Plant Protection* **2011**, *64*, 55–62. doi: [10.30843/nzpp.2011.64.5962](https://doi.org/10.30843/nzpp.2011.64.5962).

11. Goldson, S.L.; Tomasetto, F. Apparent Acquired Resistance by a Weevil to Its Parasitoid Is Influenced by Host Plant. *Frontiers in Plant Science* **2016**, *7*. doi: [10.3389/fpls.2016.01259](https://doi.org/10.3389/fpls.2016.01259).

12. Tomasetto, F.; Tylianakis, J.M.; Reale, M.; Wratten, S.; Goldson, S.L. Intensified agriculture favors evolved resistance to biological control. *Proceedings of the National Academy of Sciences* **2017**, 201618416. doi: [10.1073/pnas.1618416114](https://doi.org/10.1073/pnas.1618416114).

13. Casanovas, P.; Goldson, S.L.; Tylianakis, J.M. Asymmetry in reproduction strategies drives evolution of resistance in biological control systems. *PLOS ONE* **2018**, *13*, e0207610. doi: [10.1371/journal.pone.0207610](https://doi.org/10.1371/journal.pone.0207610).

14. Williams, C.L.; Goldson, S.L.; Baird, D.B.; Bullock, D.W. Geographical origin of an introduced insect pest, *Listronotus* *Bonariensis* (Kuschel), determined by RAPD analysis. *Heredity* **1994**, *72*, 412–419. doi: [10.1038/hdy.1994.57](https://doi.org/10.1038/hdy.1994.57).

15. Vink, C.J.; Kean, J.M. PCR gut analysis reveals that *Tenuiphantes* *Tenuis* (Araneae: Linyphiidae) is a potentially significant predator of Argentine stem weevil, *Listronotus* *Bonariensis* (Coleoptera: Curculionidae), in New Zealand pastures. *New Zealand Journal of Zoology* **2013**Taylor & Francis.

16. Kahle, D.; Wickham, H. Ggmap: Spatial Visualization with ggplot2. *The R Journal* **2013**, *5*, 144. doi: [10.32614/RJ-2013-014](https://doi.org/10.32614/RJ-2013-014).

17. Bushnell, B. *BBMap: A Fast, Accurate, Splice-Aware Aligner*; Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States), 2014;Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).

18. Chapman, J.A.; Ho, I.; Sunkara, S.; Luo, S.; Schroth, G.P.; Rokhsar, D.S. Meraculous: De Novo Genome Assembly with Short Paired-End Reads. *PLoS ONE* **2011**, *6*. doi: [10.1371/journal.pone.0023501](https://doi.org/10.1371/journal.pone.0023501).

19. Chapman, J.A.; Ho, I.Y.; Goltsman, E.; Rokhsar, D.S. Meraculous2: Fast accurate short-read assembly of large polymorphic genomes. *arXiv:1608.01031 [cs, q-bio]* **2016**. Retrieved from <http://arxiv.org/abs/1608.01031>.

20. Goltsman, E.; Ho, I.; Rokhsar, D. Meraculous-2D: Haplotype-sensitive Assembly of Highly Heterozygous genomes. *arXiv:1703.09852 [q-bio]* **2017**. Retrieved from <http://arxiv.org/abs/1703.09852>.

21. Harrop, T. HMW DNA extraction for insects. **2018**. doi: [10.17504/protocols.io.pnwdmfe](https://doi.org/10.17504/protocols.io.pnwdmfe).

22. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* **2019**, 1. doi: [10.1038/s41587-019-0072-8](https://doi.org/10.1038/s41587-019-0072-8).

23. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. doi: [10.1093/bioinformatics/btv351](https://doi.org/10.1093/bioinformatics/btv351).

24. Roach, M.J.; Schmidt, S.A.; Borneman, A.R. Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **2018**, *19*, 460. doi: [10.1186/s12859-018-2485-7](https://doi.org/10.1186/s12859-018-2485-7).

25. Smit, A.F.A.; Hubley, R. RepeatModeler Open-1.0 2015.

26. Smit, A.F.A.; Hubley, R.; Green, P. RepeatMasker Open-4.0. 2015.

27. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Molecular Ecology* **2013**, *22*, 3124–3140. doi: [10.1111/mec.12354](https://doi.org/10.1111/mec.12354).

28. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **2015**, *4*. doi: [10.1186/s13742-015-0047-8](https://doi.org/10.1186/s13742-015-0047-8).

29. Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics* **2010**, *11*, 94. doi: [10.1186/1471-2156-11-94](https://doi.org/10.1186/1471-2156-11-94).

30. R Core Team R: A Language and Environment for Statistical Computing. **2015**.

31. Lischer, H.E.L.; Excoffier, L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **2012**, *28*, 298–299. doi: [10.1093/bioinformatics/btr642](https://doi.org/10.1093/bioinformatics/btr642).

32. Foll, M.; Gaggiotti, O. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* **2008**, *180*, 977–993. doi: [10.1534/genetics.108.092221](https://doi.org/10.1534/genetics.108.092221).

33. Delaneau, O.; Marchini, J.; Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nature Methods* **2012**, *9*, 179–181. doi: [10.1038/nmeth.1785](https://doi.org/10.1038/nmeth.1785).

34. Gautier, M.; Klassmann, A.; Vitalis, R. Rehh 2.0: A reimplementation of the R package rehh to detect positive selection from haplotype structure. *Molecular Ecology Resources* **2017**, *17*, 78–90. doi: [10.1111/1755-0998.12634](https://doi.org/10.1111/1755-0998.12634).

35. Excoffier, L.; Dupanloup, I.; Huerta-Sánchez, E.; Sousa, V.C.; Foll, M. Robust Demographic Inference from Genomic and SNP Data. *PLOS Genetics* **2013**, *9*, e1003905. doi: [10.1371/journal.pgen.1003905](https://doi.org/10.1371/journal.pgen.1003905).

36. Köster, J.; Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **2012**, *28*, 2520–2522. doi: [10.1093/bioinformatics/bts480](https://doi.org/10.1093/bioinformatics/bts480).

37. Kurtzer, G.M.; Sochat, V.; Bauer, M.W. Singularity: Scientific containers for mobility of compute. *PLOS ONE* **2017**, *12*, e0177459. doi: [10.1371/journal.pone.0177459](https://doi.org/10.1371/journal.pone.0177459).

38. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE* **2011**, *6*, e19379. doi: [10.1371/journal.pone.0019379](https://doi.org/10.1371/journal.pone.0019379).

39. Sabeti, P.C.; Varilly, P.; Fry, B.; Lohmueller, J.; Hostetter, E.; Cotsapas, C.; Xie, X.; Byrne, E.H.; McCarroll, S.A.; Gaudet, R. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* **2007**, *449*, 913–918. doi: [10.1038/nature06250](https://doi.org/10.1038/nature06250).

40. Gautier, M.; Vitalis, R. Rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* **2012**, *28*, 1176–1177. doi: [10.1093/bioinformatics/bts115](https://doi.org/10.1093/bioinformatics/bts115).

41. McKenzie, J.A.; Batterham, P. The genetic, molecular and phenotypic consequences of selection for insecticide resistance. *Trends in Ecology & Evolution* **1994**, *9*, 166–169. doi: [10.1016/0169-5347(94)90079-5](https://doi.org/10.1016/0169-5347(94)90079-5).

42. Green, L.; Battlay, P.; Fournier-Level, A.; Good, R.T.; Robin, C. *Cis*- and *Trans*-acting variants contribute to survivorship in a naïve *Drosophila* *Melanogaster* population exposed to ryanoid insecticides. *Proceedings of the National Academy of Sciences* **2019**, 201821713. doi: [10.1073/pnas.1821713116](https://doi.org/10.1073/pnas.1821713116).