final report

Project code: B.AHE.0315

Prepared by: Peter Hunt
Commonwealth Scientific and Industrial Research Organization

Date published: 18 January 2019

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

New option for monitoring drench resistance and movement of Barber's Pole Worm

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Executive summary

The continued importance of internal parasites' negative impact on profitable livestock production, complicated by chemical and drug resistance, led MLA, in the 2016 annual call for project proposals, to publish a terms of reference (ToR) document "Reduce the economic impact of endemic diseases of sheep and grassfed cattle". The ToR included three specific areas to be addressed. Two of these were addressed by our proposal (B) "modelling the potential southward movement of livestock diseases in response to climate change and the adaptation in a future whole farm systems context," and (C) "reducing the impact of major endemic diseases in the national sheep flock, particularly internal parasites". This project aimed to develop a DNA-based test which could be used by producers to monitor drench resistance and parasite management practices within livestock enterprises. The adoption and use of such a test by producers would also accumulate data which subsequently would be informative at the industry level for monitoring the geographic spread of species of importance, and in particular strains of species which have deleterious characteristics such as drug resistance. We developed the test for Barber's Pole Worm (BPW), as it is an important parasite of sheep, and because we had preliminary data available for this species which helped us complete the work over a shorter timeframe than might be necessary for other species.

The use of anthelmintic drugs, or drenches, has been a major management intervention used for more than 50 years. Over this time, drench resistance has arisen in all species of importance in Australian sheep. Unfortunately, the release of new drenches containing new active compounds has declined, with only two new chemicals with novel modes of action entering the sheep drench market in the past four decades. The use of these drenches inevitably leads to selection for resistance, leaving only resistant individuals to breed and produce the next generations of the parasites. If a proportion of the parasites is not exposed to drench selection, then these can mix with selected individuals, slowing the progress of resistance selection for the population as a whole. Gastro-intestinal nematodes (GIN) typically exist in large numbers upon pasture and also within classes of stock which do not need regular drenching. These non-selected populations (refugia) can slow down the development of resistance for the population as a whole. Selection intensity (drenching frequency and dosage) and managing refugia are therefore important components which influence selection pressure (speed of selection) and the population genetics of parasites on farms. A third component, migration, is a factor which can significantly alter the outcomes of selection pressure so that they differ from the outcomes predicted from treatment frequency and refugia levels alone. The effect of introduced worms upon the development of drench resistance is a major knowledge gap. These two factors, selection pressure and migration are central components of the discipline known as population genetics. This project has developed a tool for understanding the population genetics of BPW. Future projects could develop similar tools for other disease organisms, including the other GIN important for sheep production.

This tool can be used by producers to monitor BPW drench resistance in sheep. We have called the tool a "genetic drive chip", and it is a DNA-based test which evaluates genetic diversity in BPW populations. The tool can detect changes in genetic diversity over time due to chemical selection pressure and the mixing of populations due to migration. Estimating both of these is important for interpreting the outcome of the test. We envisage the test being used by producers annually as a replacement for a faecal egg count reduction test (FECRT; a method for determining the drench susceptibility of the GIN in a sheep flock), at a fraction of the cost, and with no additional input of on-farm labour. The low cost has been achieved after extensive screening of the molecular markers used in the "genetic drive chip" to ensure they are suitable for analysing a single DNA sample produced from a larval culture containing a pool of thousands of parasites, rather than the far more costly alternative of analysing individual worms. The "genetic drive chip" for BPW uses single nucleotide polymorphisms (SNPs; which are also used to estimate the genomic breeding values of

livestock) in combination with management data from the property. The "genetic drive chip" uses records of farm management data such as stock movements, stock purchases and drench usage with the DNA-test to predict drench resistance and monitor over-all BPW control performance on the farm.

This project demonstrated that the test can detect the difference between worm populations and can detect when two populations have been mixed together, giving us the ability to monitor worm migration. The genetic drive chip can be used to detect the increase in drench resistance over time in concert with good farm management records. In two experiments, one using the drugs closantel, abamectin and oxfendazole, and the other using monepantel, and it was shown that particular SNPs are important for monitoring the drenches used. Different SNPs are needed for different drenches, which differ in their modes of action.

Six key messages emerge from this work:

1- The BPW genetic drive chip is a new drug resistance diagnostic method, which utilises a panel of over 100 SNP assays.
2- The project has shown that the complexity of drench resistance genetics does not allow simple, single gene tests for resistance diagnostics.
3- The BPW genetic drive chip is likely to be less expensive than FECRT, with a substantial decrease in labour needed on farm.
4- Further field evaluation and supply chain testing project, over three years and in collaboration with at least 100 producers, is recommended before the BPW genetic drive chip can be commercialised.
5- The genetic drive chip can be further developed to incorporate additional worm species of sheep and cattle, with very little increase in the retail price of the diagnostic test.
6- A commercial partner, licensed to produce, distribute and administer the test, will have to be found.

Table of contents

1 Background

1.1 Drug-resistance in Barber's Pole Worm

1.1.1 Barber's Pole Worm in Australia

In 1967, MacKenzie published a map of the distribution of Barber's pole worm (BPW or *Haemonchus contortus*). The map showed a BPW endemic zone along the tablelands and western slopes from central NSW to central Queensland. Since that time, geographic prediction of endemic areas based on climate, latitude and altitude has progressed. We can now say with confidence for example, that sheep and goats kept on the northern coastal areas of NSW and the Southern coastal areas of Queensland will be exposed to BPW, even though these areas are not included on the 1967 map. Further changes in our understanding of the extent of the BPW endemic and sporadic zones have been aided by reports from diagnostics laboratories over the past two decades. Sporadic occurrences of BPW now occur over many additional areas in all states. Multiple factors have contributed to these observed changes, notably an increase in reporting, but also changes to stocking densities, climate and the distribution of sheep and goat enterprises. The distribution of feral goats and deer may also have contributed. Simultaneously, an increase in both the diversity of chemical control measures and the number of non-chemical control options has made parasite control more sophisticated, but also more complex. In the face of a changing parasite risk landscape, producers need better tools to help make crucial decisions.

In Australia, drench resistance in BPW was first identified in 1966 in a population in NSW that showed increased resistance to thiabendazole (Love 2011). Since then, BPW populations have evolved resistance to almost all drench actives (Fig 1). BPW populations in large areas of NSW and QLD are estimated to have 90% resistant to the benzimidazole (BZ) and 40-60% resistance to levamisole (LEV). An estimated 10% of BPW populations in the northern NSW and southern QLD region are estimated to be multiple resistant to three drench compound groups with separate modes of action; BZ, LEV and macrocyclic lactones (ML) (Emery *et al* 2016, Love 2011, Kotze and Prichard 2016, Lyndal-Murphy *et al* 2014).

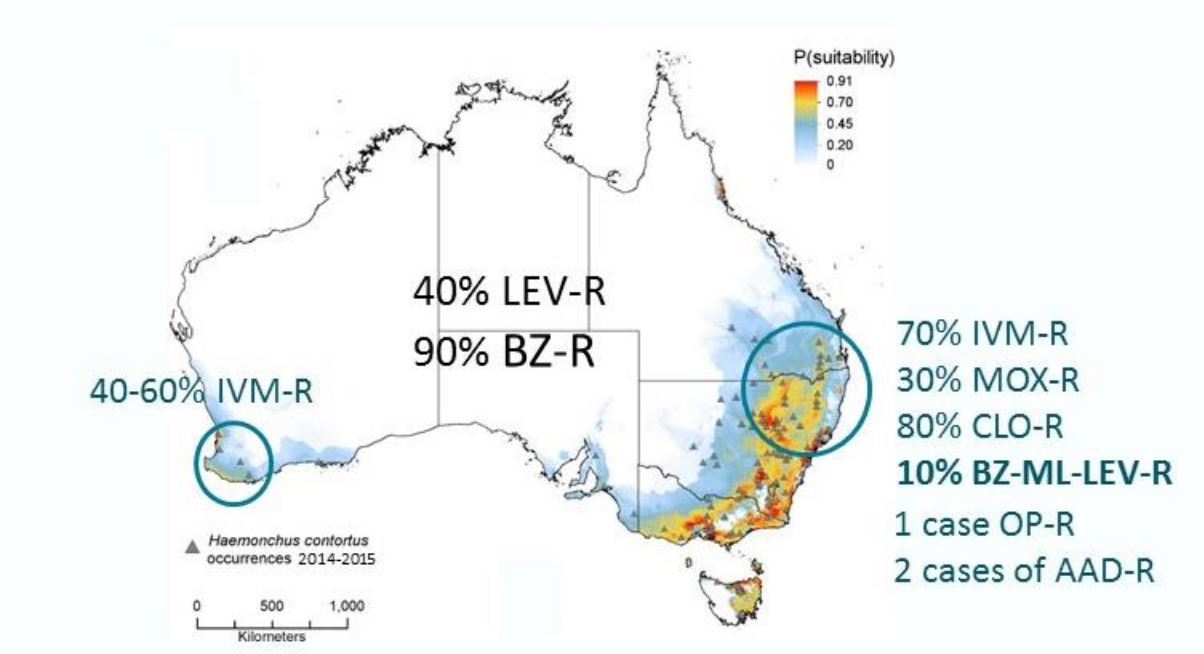

Fig 1. Drench resistance levels in BPW populations as of September 2018. Original map of BPW occurrence from Emery et al 2016.

Drench resistance, just as other examples of pesticide resistance in Australia, is increasing every year and endangering food and fibre production through animal health and production costs. BPW is the most pathogenic gastro-intestinal nematode (GIN) parasite which effects the Australian sheep production industry and is responsible for a large part of the $436 million in production losses in Australia (Lane *et al.* 2015). A conservative cost of parasitism has been found to be $12h/sheep (Lane *et al* 2015)**.**

The ParaBoss suite of web-based information has been a key development, but the advice there is necessarily not enterprise- or season-specific. Better management tools will enable producers to make the most of online advisory services. We aim to expand the availability of useful tools that producers can access. No option currently exists for detecting an incursion of any parasite species onto a property where other strains of the parasite already occur, and the options for predicting drench resistance are expensive and laborious.

1.1.2 GIN drug resistance diagnostics

Currently, the faecal egg count reduction test (FECRT), a labour intensive (30+ days) and costly test (between $600-800), is the only option for measuring drench resistance. The majority of sheep producers are aware and worry about the threat of drench resistance (Aleta Knowles, Virbac, ASP 2018), however, only an estimated 1% use the FECRT diagnostic. A lack of knowledge, in particular a low acceptance of the fact that resistance is inevitable contributes to this low adoption rate, however the high amount of labour required to conduct FECRT, which increases with every drug tested, and the high price of the test exacerbate this problem.

DNA tests for resistance offer the livestock producer a far simpler and more cost effective alternative to FECRT (Fig 2). They have not been developed to this point, because the number of actives and different species of parasites made their development expensive and complicated in the past. At the beginning of this project we were in a position to test a completely new type of DNA test (Genetic Drive Chip) which could work for any active because it did not rely on specific information about genes associated with resistance. We proposed to evaluate the idea over two years, using BPW as our target nematode species.

.

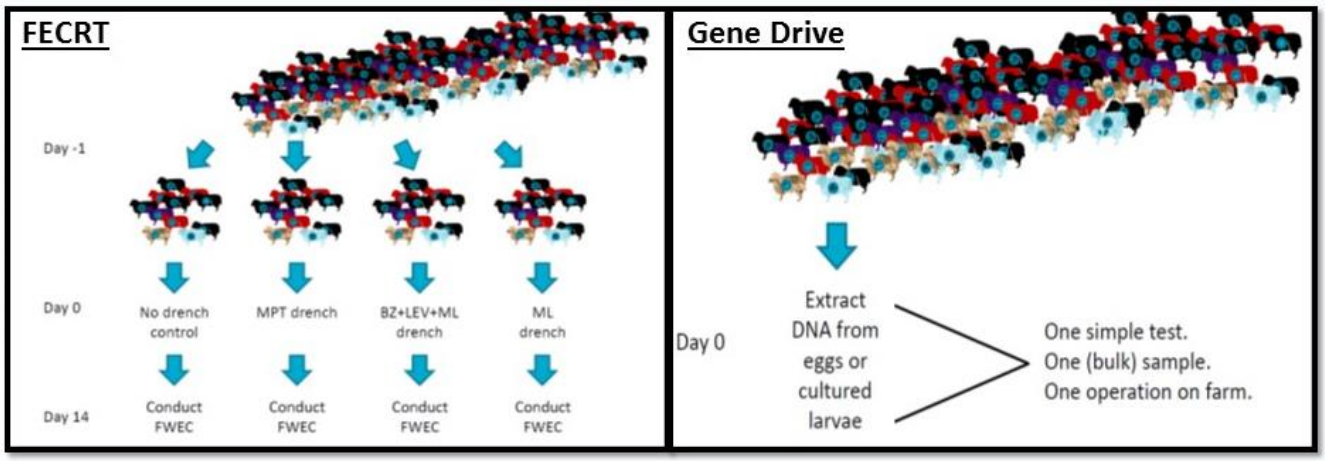

Fig 2. Comparison of the FECRT and Gene Drive Chip method in time and labour needed.

1.2 Genetic Drive Chip overarching aims

1. In preparation for delivering a DNA-based test (the Genetic Drive Chip) for producers to use in place of FECRT

a. Develop a set of genetic markers (SNP platform) for monitoring BPW on farm

b. Demonstrate that the SNP platform can detect population changes which are indicative of drug resistance, or the introduction of parasites onto a property

2. Provide two years of postdoctoral training for a new scientist within Australian livestock parasitology, an essential capability which needs support to be sustainable.

At the successful completion of these studies, subsequent work to develop the Genetic Drive Chip for use within sheep producer's enterprises will be able to proceed with confidence. In addition, a new scientist will have been trained who will be able to work at the interface between molecular diagnostics and on-farm parasite control.

1.3 Significance for the livestock industry

The outcome of the work is to provide a tool for producers to use in parasite management which can detect important parasite population changes on farm such as the development of drench resistance or a failure of on-farm quarantine. The information provided to the producer when using the Genetic Drive Chip, can be used to guide strategic decisions regarding parasite management and the use of drenches. The tool provides valuable feedback to the producer regarding the effectiveness of on-farm quarantine measures, the effectiveness of drenches, the mix of parasite species present on their property, and the subsequent threat to sheep production. The Genetic Drive Chip could improve livestock health, lower input costs and labour and ensure that only effective drenches are used, and that they are only administered when appropriate to reduce ongoing selection for drench resistance. The Genetic Drive Chip could also monitor the effectiveness of current and future worm management strategies for the whole industry.

2 Project objectives

2.1 Genetic Drive Chip

The project aimed to develop a tool, the Genetic Drive Chip, using large collections of DNA-based neutral markers in a high throughput format to assess changes within, and differences between, *H. contortus* populations. This would enable determination of new incursions of Barber's Pole Worm (BPW), and also whether changes have occurred within a farm (e.g. because of drug resistance).

By 1st November 2018, three objectives were to be met.

1) Select Single Nucleotide Polymorphism (SNP) and insertion/deletion (InDel) markers from appropriate genome regions for the chip.

2) Use the chip to analyse samples from five geographically separated BPW isolates with different chemical resistance profiles, to distinguish populations from each other.

3) Use the chip to track genetic change following chemical selection pressure

3 Methodology

3.1 Single nucleotide polymorphism discovery

Prior to the start of this project, DNA was extracted from five and six pools of 20 BPW individuals from the isolates Wallangra2003 and McMaster1931 respectively. A total of 11 paired-end 100 bp fragment libraries were sequenced using MiSeq technology (Illumina Australia, VIC). The variant discovery pipeline (Fig. 3, Appendix 9.1) consisted of individual library alignment to the two available BPW reference genome assemblies (Laing *et al.* 2013; Schwarz *et al.* 2013).

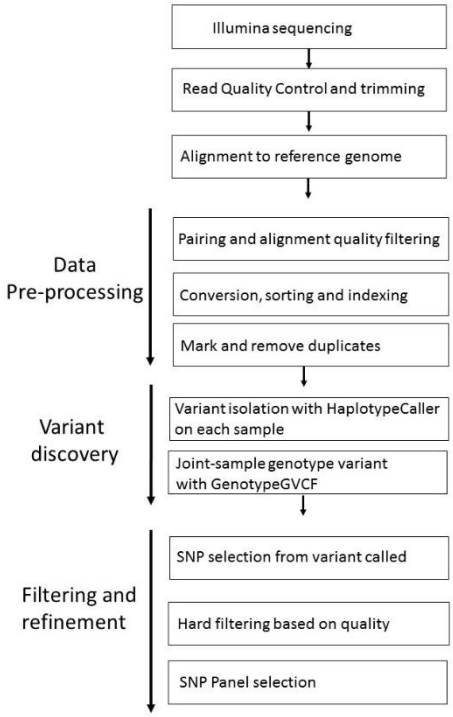

Figure 3. Workflow for SNP marker discovery in *H. contortus* (BPW) using 5 pooled samples of Wallangra2003 isolate and 6 pooled samples of McMaster1931.

Following the alignment, reads were filtered based on high alignment quality (MAPQ > 60), also ensuring that both reads in a pair were aligned using samtools (v1.3.1, Li *et al.* 2009). The resulting SAM files were converted to BAM files, sorted by reference genome coordinate and indexed. Optical and PCR read duplicates were identified and removed using Picard (v2.9.2,

http://broadinstitute.github.io/picard). The variant discovery task consisted of two steps. Variants (SNPs and InDels) were identified in each of the 11 samples individually using GATK HaplotypeCaller tool (Van der Auwera *et al.* 2013), then the two isolates were genotyped using GATK GenotypeGVCF tool.

In the first step of the filtering and refinement task, the resulting vcf file was filtered to discard InDels and keep only SNPs. Following GATK best practice for variant calling (DePristo *et al.* 2011), a hard filtering step was applied to eliminate low quality SNPs. The hard filter consisted of eliminating variants with a low quality by coverage score (QD<5), variants found on the same read orientation (FisherStrand >60), variants which have the reference and alternate variant call of different quality (MQRankSumtest<-5), variants that were found only within the extremities of reads (ReadPosRankSum <-8), and finally variants found more in one strand (reverse or forward) than the other were also eliminated (StrandOddRatio>3).

The SNP panel selection was performed. A total of 250 selected SNPs were chosen across the remaining high quality SNPs identified following several criteria; equal number of transition and transversion substitutions, selection across the largest 50% of all genomic scaffolds, across the range of possible allele frequency differences between populations (0-1) and across the range of possible minor allele frequencies within populations (0-0.5). The rationale is that these markers have unknown association with drug selection, the majority being neutral and, when analysed as a pool, will reflect the overall genetic changes in the population under study. A further 84 SNPs were selected specifically to enable comparisons with SNPs in genes implicated in resistance from the scientific literature. To achieve this, protein sequences from genes with known and putative effects on drug susceptibility phenotype in *H. contortus* (Appendix 9.2) were retrieved from NCBI, and screened against the list of high quality SNPs discovered. These included known SNPs which are putatively involved in drug resistance, and were from the isotype 1 of the β-tubulin gene (BZ resistance), monepantel target gene *acr* (MPL resistance), the glutamate gated chloride ion channel-3 (MPL resistance), and the amphid dendrite dye-filling deficient gene *dyf-7* (ML resistance). Further, putative genes involved in drug resistance such as cytochrome monooxygenase P450s, P-glycoprotein, UDP-glucuronosyltransferase, nicotine acetylcholine receptors, ATP-binding cassette transporters, haf-transporters, and other glutamate gated chloride ion channel receptors were also screened for presence of high quality SNPs (Appendix 9.2). A further criteria for selecting SNP was to ensure the particular SNP is the only variant at 150 bp upstream and downstream in the genome, to assist in the creation of the allelotyping reactions.

In late 2018, a new genome assembly for *H. contortus* was released on WormBase from the Wellcome Trust Sanger Institute (Assembly HCON_V4, GCA_000469685.2) assembling the genome into six chromosomes with a total of 283,439,308 bp for a total of 19,430 coding genes (WormBase version WBPS 12). The SNP panel was screened against this new assembly to assess SNP position and functionality after most of the work had been completed in the project.

3.2 Allelotyping reaction

The 334 biallelic SNPs selected were used to design multiplex assays to be genotyped using the Sequenom platform. An average of 25-40 SNPs were included in each reaction ("plex" see Appendix 9.2). The allelotyping is performed by analysing primer extension reactions where the added bases

has a mass marker, enabling differentiation between the two allele bases by mass spectrometry. Polymerase chain reaction (PCR) primers and extension primers were designed for each SNP. The markers were multiplexed in as few sets as possible depending on the mass of the allele sequenced and ensuring each marker's amplicon size did not overlap on the chromatogram. Each SNP marker was analysed after primer extension using the Agena Bioscience Sequenom MassArray system in triplicate (Neogen Australasia [formely UQ Animal Genetics Laboratory], Gatton, QLD). The MassArray technology has not been robustly tested on pooled eukaryotes samples, and had not been tested for *H. contortus* pools previously. **The sequences of primers used for amplification and primer extension and the sets which go together to form each "plex" is sensitive intellectual property which cannot be protected except as a "trade secret".**

3.3 Genetic Drive Chip quality control

The quality control for each SNP involved three steps: the sample pooling versus individual allele frequency estimation, the successful amplification in co-occurring species and the identification of admixed populations. The quality control pipeline is illustrated in Fig 4 and the R script is available in Appendix 9.3.

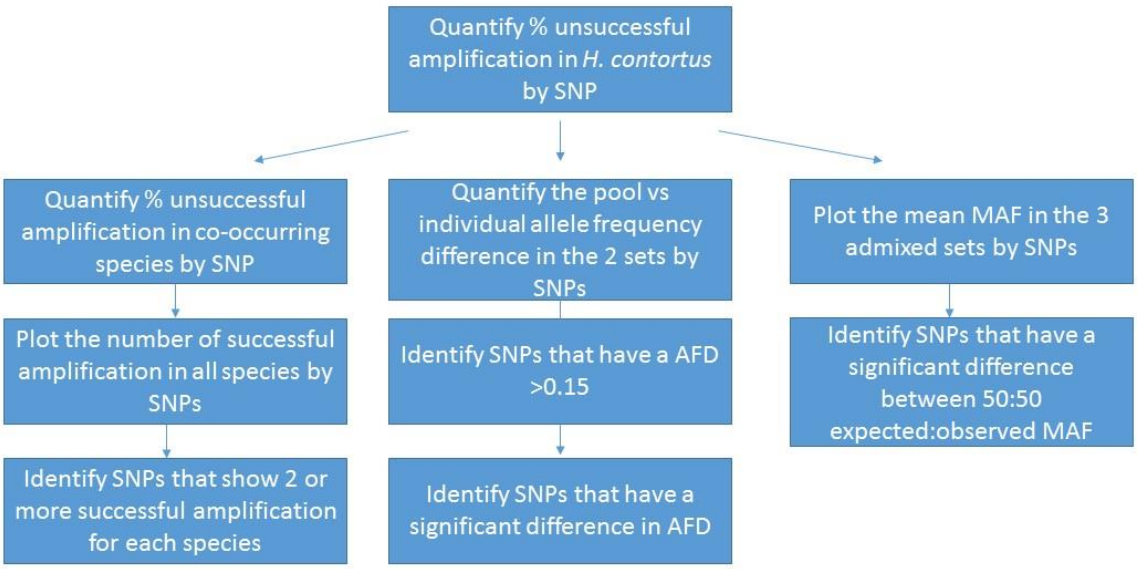

Fig 4. Genetic drive chip quality control pipeline. Quality control script available in Appendix 9.3.

3.3.1 Comparing allelotyping for individuals compared to pools

DNA from 30 adult male individuals from the McMaster1931 and the Wallangra2003 isolates was extracted. DNA was quantified using spectrophotometer (NanoDrop, Thermo Fisher Scientific) and the 30 individuals per population were pooled in equimolar concentration to create a MCM30 and WAL30 pool sample. These two pools were then also pooled in equimolar concentration to create a TOT60 sample. Each of the 60 individuals and three pooled samples were genotyped for each SNP marker.

Pooled individual samples were used for subsequent assessment of populations of *H. contortus* by Sequenom allelotyping. This method was selected because 1) field derived samples of *H. contortus*

can only be obtained from nematode eggs in faeces or larvae from faecal cultures, larvae are preferred because of a lower risk of PCR-inhibiting contaminants, 2) *H. contortus* larvae are <1 mm long, therefore time consuming to separate and do not reliably yield DNA, and finally 3) in order to obtain a reliable estimate of population genetics statistics in populations that contain millions of individuals, several hundred to a thousand individuals would be needed. Pooled genotyping has recently been acknowledged to be a cost-effective and reliable estimate of individual data for population genetics applications (Futschik and Schlötterer 2010; Boitard *et al.* 2012; Rellstab *et al.* 2013; Lynch *et al.* 2014).

3.3.2 Detection of mixed population

One of the objectives of the Genetic Drive Chip is to detect the presence of gene flow from one population into another such as when infected sheep are introduced to a property from elsewhere. Therefore we artificially created admixed populations using three population pairs; 1) the MCM30 and WAL30, 2) Mackay and Cannawigara, and 3) Goondiwindi and GoldCoast (see details of populations in table 2). These three pairwise sets were selected from the samples available because the WAL/MCM pair was initially used in genomic discovery of the SNPs, the Mackay/Cannawigara pair were from the most divergent North/South, and the Goondiwindi/GoldCoast pair were from the most divergent Coast/Inland distance. Pooled samples from each population were mixed at different ratios in a pairwise fashion; 0:100, 25:75, 50:50, 75:25, and 100:0. The alternate (e.g non-reference) allele frequencies from these admixed populations were evaluated in two ways. First, the overall observed vs expected increasing or decreasing frequencies were evaluated. Second, the alternate allele frequency (ALT) observed for the 50:50 population was compared to the observed mean of the 0:100 and 100:0 population allele frequencies.

3.3.3 Co-occurring parasitic nematode species

In natural conditions, *H. contortus* usually occurs in a community of GIN parasites. The GIN most commonly found in Australia alongside *H. contortus* are *Trichostrongylus vitrinus*, *Trichostrongylus colubriformis*, and *Teladorsagia circumcincta, with Oesophagostomum venulosum*, *Chabertia ovina*, *Cooperia oncophora* and *Oesophagostomum columbianum* as the next most common. It was therefore important to ensure the *H. contortus*-specificity of our SNP panel so that the test would be useful to analyse samples from co-infected sheep. Adult individual samples from *O. venulosum, C. oncophora* and *T. vitrinus* were not available in-house; *O. venulosum* was acquired from the South Australian Museum and from collaborators at the University of Melbourne (Table 1). Although not available as a single species sample, a 20:80 sample of *H. contortus*:*T. vitrinus* was available in house and was used to establish SNPs affected by this level of *T. vitrinus* inclusion in *H. contortus* samples. A sheep (*Ovis aries*) DNA sample was also used to ensure no cross-reaction with host genetic material.

Table 1. Co-occurring species samples used against the BPW genetic drive chip.

Species (common name)	Location and Date collected	Sample original location
Trichostrongylus vitrinus (black scour worm) Contaminated with 20% *H. contortus*	Armidale 18/05/2007	CSIRO McMaster collection
Trichostrongylus colubriformis (black scour worm)	Armidale 05/09/2012	CSIRO McMaster collection
Teladorsagia circumcincta (brown stomach worm)	Armidale 28/06/2007	CSIRO McMaster collection
Oesophagostomum venulosum (large bowel worm)	Unknown location in Australia 01/03/1981	South Australian Museum
Chabertia ovina (large mouthed bowel worm)	Armidale 28/06/2007	CSIRO McMaster collection
Cooperia oncophora	Armidale 09/02/2013	CSIRO McMaster collection
Oesophagostomum columbianus (nodule worm)	Armidale 28/06/2007	CSIRO McMaster collection
Ovis aries (sheep)	Armidale 2016	NA

3.4 Spatial analysis within Australian BPW populations

A total of 34 BPW populations from livestock across Australia and South Africa were available. Ten populations from commercial properties were sampled in 2017/18; populations including those from Holbrook, Goulbourn, Sydney, Tenterden, Duri, Bombala, Cooma, Wongarbon and Wellington came from Dawbuts; anonymous populations 2, 4, 10, 13, 14, 15, 2140, 2141, 2142, 2143 came from Invetus. Twenty four historical samples came from CSIRO livestock parasite collection at Armidale (Table 2). The two samples from WA, were obtained from DAFWA, via Dieter Palmer (Harvey) and Jill Lyon (Narrikup). The two South African samples were included as outgroups to the Australian populations, as well as to evaluate the use of the SNP panel in international *H. contortus* isolates.

Table 2. Field populations sampled to assess population genetic differentiation using the BPW genetic drive chip. BZ; benzimidazole, ML; macrocyclic lactone, LEV; levamisole, SAL-P; salicylanilide phenol, MPTL; monepantel, IVM; ivermectine, ABA; abamectin, CLO; closantel, OP; organophosphate, R; resistant, S; susceptible. Lab-derived populations refer to populations that have been either been selected artificially in vivo for a specific phenotype or populations that have been inbred for at least 10 years in experimental conditions.

Population	Collection year	State	Resistance phenotype	Origin
McMaster1931	2006*	NSW	Susceptible to BZ, ML, LEV, SAL-P, MPTL and OP	Lab-derived
Wallangra2003	2006*	NSW	R to BZ, ML, LEV and SAL-P	Lab-derived
Bundarra	2014	NSW	NA	field
Mackay	2009	QLD	NA	field
GoldCoast	2004	QLD	R to ML	field
Cannawigara	2006	SA	Low R to BZ	field
Goondiwindi	2011	QLD	NA	field

ChiswickAVRS	1999	NSW	ML resistance and smooth vulvar phenotype	Lab-derived
Narrikup	2017	WA	NA	field
Harvey	2017	WA	NA	field
Kirby	2006	NSW	Susceptible to all drugs	field
Guyra	2007	NSW	NA	field
Onderstepoort	1990	South Africa	Susceptible to all drugs	Lab-derived
White River	1990	South Africa	R to IVM, BZ, and SAL-P	Lab-derived
Holbrook	2018	NSW	NA	field
Goulburn	2018	NSW	38 days post moxidectine (ML) and treatment	field
Sydney (Campbelltown)		NSW	40 days post ABA, oxfendazole and LEV treatment	field
Tenterden	2017	NSW	2 months post monepantel treatment	field
Duri	2017	NSW	NA	field
Riverina (Tumblong)	2017	NSW	No drench used	field
Bombala		NSW	2 months post ABA and LEV and BZ (albendazole) treatment	field
Cooma	2018	NSW	3 ½ months post ABA and CLO treatment	field
Wongarbon	2017	NSW	7 months post unspecified drench treatment	field
Wellington	2017	NSW	2 months post ABA and CLO treatment	field
2	2018	NSW	NA	field
4	2018	NSW	NA	field
10	2018	NSW	NA	field
13	2018	NSW	NA	field
14	2018	NSW	NA	field
15	2018	NSW	NA	field
2140	2018	NSW	NA	field
2141	2018	NSW	NA	field
2142	2018	NSW	NA	field
2143	2018	NSW	NA	field

* population passed through sheep without any drug selection.

3.4.1 Population genetics differences

Population-level SNP allele frequencies were obtained from each of these samples. The mean frequency of the non-reference alternate (ALT) allele of each SNP for each population was calculated along with the respective mean frequency uncertainty from the Sequenom allelotyping. The mean frequency uncertainty from allelotyping was added to any frequency difference between individual and pool to get an overall frequency uncertainty value. These overall frequency uncertainty values were used to create a distribution of potential ALT frequencies for each population (Fig 5). The R script used for these analysis is available in Appendix 9.4.

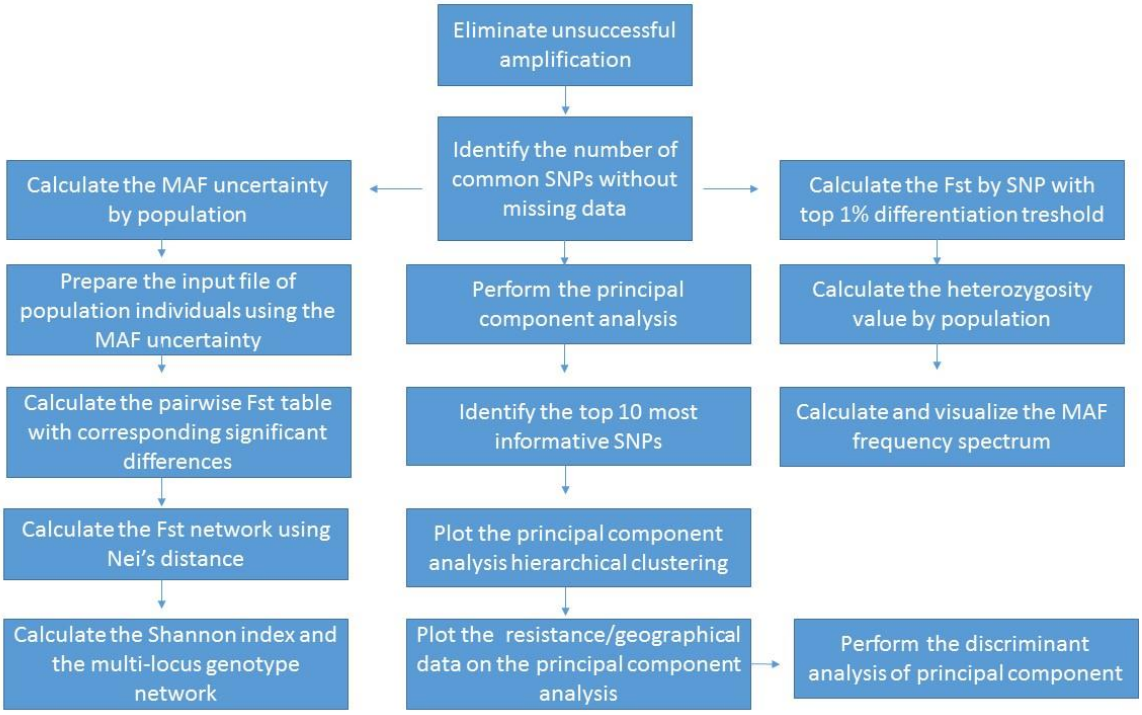

Fig 5. Population genetics analysis pipeline. Complete R script available in Appendix 9.4.

A series of population genetics and visualization methods were used to analyse changes in allele frequencies. Principal component analysis (PCA), pairwise fixation index (Fst), hierarchical clustering and a pairwise Fst distance neighbour-joining tree were used to validate the relationships between populations. The neighbour-joining tree was created by first calculating the distance between each Fst pairwise population pair using ALT frequency, the unweighted algorithm was used to compute dichotomous clades, the tree was transformed into an equal angle network, and finally, 500 bootstrap permutations were computed. The Fst statistic was also calculated between SNPs (e.g. Fst by loci) in order to identify markers with signature of selection pressure over the baseline standard variation in Fst values.

Additionally, population heterozygosity values were obtained for each population by calculating the mean minor allele frequency value across all common SNPs. The Shannon diversity index was also calculated for each population using the ALT frequency. The Shannon diversity index (Shannon and Weaver 1949) will report a higher diversity if there is a large variation of ALT frequencies (from 0-1) across all SNPs in the given population. Finally, a multi-locus genotype network was created by calculating the number of shared genotype (same ALT frequency within 0.05 range) across all common SNPs. Statistical analyses were conducted in R (v.3.4.0, R Development Core team 2017).

3.5 In vivo drug selection trial

3.5.1 Infection cycle

Four parental isolates were selected for their known resistant/susceptibility phenotypes for specific anthelmintic drugs:

- Wallangra2003ABACLO is resistant to macrocyclic lactones (active ingredient abamectin - ABA) and salicylanilides (active closantel - CLO)
- Wallangra2003OXF is resistant to benzimidazoles (active oxfendazole - OXF)
- McMaster1931 is susceptible to all anthelmintic active ingredients
- Bundarra2014 is resistant to monepantel (MPL) and other actives.

The Wallangra2003 isolate was originally collected in 2000 and found to be resistant to multiple drug classes including macrocyclic lactones, benzimidazoles, salicylanilides and tetrahydropyrimidines but was susceptible to the aminoacetonitrile derivate MPL (Love *et al.* 2003). Field collected Wallangra2003 was passed through sheep some of which were treated with OBZ and some treated with ABA and CLO, therefore creating new field-derived isolates with specific increased resistance to these drugs. The two new populations, Wallangra2003ABACLO and Wallangra2003OBZ, were then cryopreserved. The McMaster1931 isolate was collected from the field in 1931, was passed through sheep hosts continually without any selection pressure until 1980 when it was cryopreserved. The Bundarra2014 isolate originated from goats multiply treated with MPTL (4mL of 25 g/L), OBZ (2mL of 45.3 g/L) and ABA (5mL of 0.8 g/L), where 21 of 54 remained parasitised thirteen days following treatment (part of MLA project P.PSH.0672). The larvae cultured from samples obtained from these animals were used to infect a sheep, which was subsequently treated with MPL and remained parasitised post treatment. Larvae cultured from faecal samples obtained from this sheep were used in our experiments. In our experiment, two new lines were created by mixing either Bundarra2014 or the two Wallangra2003 variants, with the susceptible McMaster1931 isolate (Fig. 6).

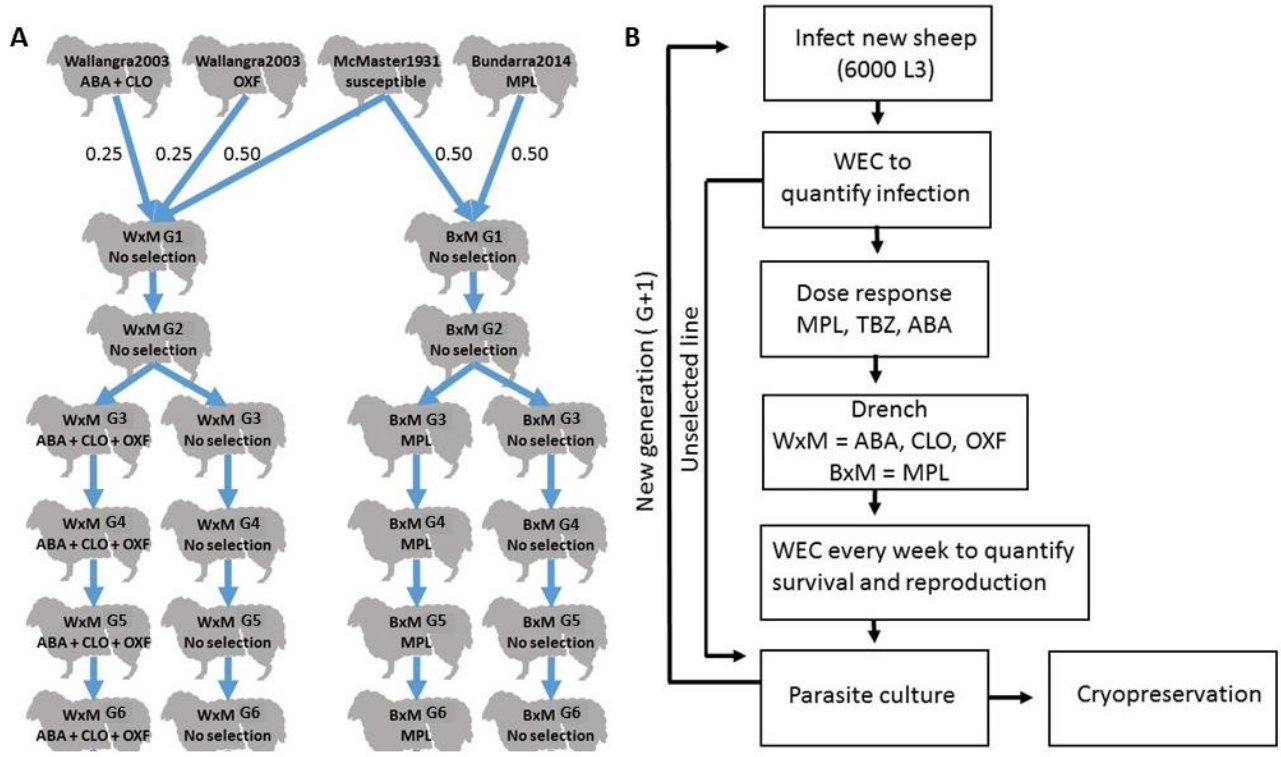

Fig 6. Drug selection trial methodology. (A) Mixed isolate lines and specific drug selection over five passages. (B) Protocol used for each passage.

The mixed isolate lines were used to infect 6 sheep (G1 in Fig. 6A), and the experiment proceeded, keeping these lines separate by culturing larvae separately for each animal and infecting a subsequent sheep for five cycles (Fig. 6B). No drug selection was applied during either the G1 generation or the subsequent G2 generation. This was done in order to allow individuals within each line to have opportunities for mixing and inter-mating without the influence of drug selection.

Eggs collected from faecal samples of sheep containing G2 adults were cultured to produce infective larvae and 6000 L3 were used to infect two new sheep per line, doubling the number of lines at this point. From G3 onwards, three lines were treated with anthelmintic drugs while the other three sheep were not treated. The three replicate Bundarra2014 X McMaster1931 (BxM) lines were selected with one dose of 25 g/L monepantel (Zolvix, 25 mg/mL monepantel, Elanco Australia) according to the manufacturer's recommended dose after the initial infection was quantified by FWEC (Fig. 3B). The three replicate Wallangra2003 x McMaster1931 (WxM) lines were selected with one dose of 45.3 g/L of oxfendazole and 120 g/L of triclabendazole (Flukazole, Virbac Australia) according to the manufacturer's recommended dose, as well as 1 g/L of abamectin and 50 g/L of closantel (Avomec Dual, Merial Australia) according to the manufacturer's recommended dose, after the initial infection was quantified by FWEC. All host sheep were treated with 0.5 mL of dexamethasone trimethylacetate (Trimedexyl, Ilium Australia) to suppress immunity and ensure that these populations were large enough for selection to proceed without losing H. contortus lines. Faecal cultures were conducted at each generation to produce infective L3 for the subsequent infection and also to cryopreserve larvae as a resource for DNA analysis and an insurance against the generation being lost. Once sufficient larvae had been cultured, animals were euthanized to obtain adult nematodes for DNA analysis.

3.5.2 Phenotypic increase in drug resistance

The susceptibility of the drug-selected BxM and WxM lines was evaluated in parental isolates and at each generation, while unselected lines were evaluated at G3 and G6 only. Both lines were evaluated against four anthelmintic classes; macrocyclic lactones using abamectin, benzimidazoles using thiabendazole, amino acetonitrile derivates using monepantel, and imidazothiazoles using levamisole. The susceptibility to each anthelmintic was evaluated in a larval development assay (LDA) as described in Kotze *et al.* (2009). The LDA has been shown to be effective at quantifying resistance to each of these chemical classes with *H. contortus* (Lacey et al 1990; Gill et al 1995; Raza et al 2016, Ruffell et al 2018). Parasite eggs were recovered from faeces by filtration using metal sieves, and further separated from debris via centrifugation on a sucrose gradient. Stock solutions of each drug were prepared in DMSO at 10 mg/mL. A series of two-fold dilutions were then prepared in order to provide the assay concentration ranges for each drug as shown in Table 3. The drugs were added to the wells of 96-well microtitre plates, and overlayed with agar. A total of 80 nematode eggs were added to each well and incubated overnight at 27°C. Growth media containing live *Escherichia coli* was then added to each well, and development was allowed to continue for a further six days. Development from egg to L3 was then assessed by counting the number of fully grown L3 larvae in each well of the assay plate. Commercial product Zolvix was used as a source of monepantel, technical grade abamectin was used (ChemService Inc., West Chester, USA), and technical grade thiabendazole and levamisole were used (Sigma-Aldrich, Australia). Each dose of

each drug was triplicated and each plate contained six control wells. Replicates were analysed separately in G1, G2, and G6 and pools of BxM and WxM were analysed in G3, G4, and G5.

Table 3. Anthelmintic drug and doses used in larval development drug-response assays in *H. contortus* lines.

Drug class	Active ingredient	Drench	Dose range
Amino acetonitrile	monepantel	Zolvix	0.000397-26.0000
Benzimidazole	thiabendazole	Technical	0.00127-20.8
Macrocyclic lactone	abamectin	Technical	0.0000398-0.0203
Tetrahydropyrimidine	levamisole	Technical	0.00127-10.4

The number of fully grown L3 in each well was expressed as a percentage of the mean number of L3 in multiple control wells. This data was analysed using non-linear regression with GraphPad Prism ® software (GraphPad Software Inc., USA, version 6.01). If the raw data showed a biphasic response, then the data was separated at the plateau and data was reanalysed separately. A one-way ANOVA was performed using EC50, standard error and degrees of freedom, followed by Tukey's HSD test to identify the significant differences in EC50 between pairwise populations. Analysis was performed in GraphPad prism.

3.5.3 Allele frequency changes following drug selection

Generation differentiation was visualized using the principal component analysis (PCA), hierarchical clustering, Fst by loci and pairwise Fst comparison, ALT frequency spectrum, and heterozygosity value comparison. The same methods were used as described in spatial dataset analysis above. The same population genetic script as above was used (Fig 5).

3.5.4 Early prediction of drug resistance

Using the longitudinal dataset, we used corresponding generation by treatment by SNP locus population-wide minor allele frequency to which we added a column for each of the 4 drugs tested using the LDA (using the anthelmintics levamisole, oxfendazole, abamectin and monepantel). In this drug-specific column we added the EC50 value for each generation by treatment combination. An additional column using categorical variables (e.g. susceptible or resistant) was also used to assess the ability to predict increased resistance to drugs. This input file was processed using the Random Forest algorithm (Breiman 2001), which is a machine learning method that relies on estimating the number of decision trees needed to predict a state from the other variables from multiple permutations (Fawagreh *et al* 2014). It also allows assessment of the variable interaction in prediction accuracy, the importance of each variable, and prediction of the value of the independent variables that best predict the status (Fig 7). Data were analysed in R using package RandomForest (script available in the population genetics script in Appendix 9.4).

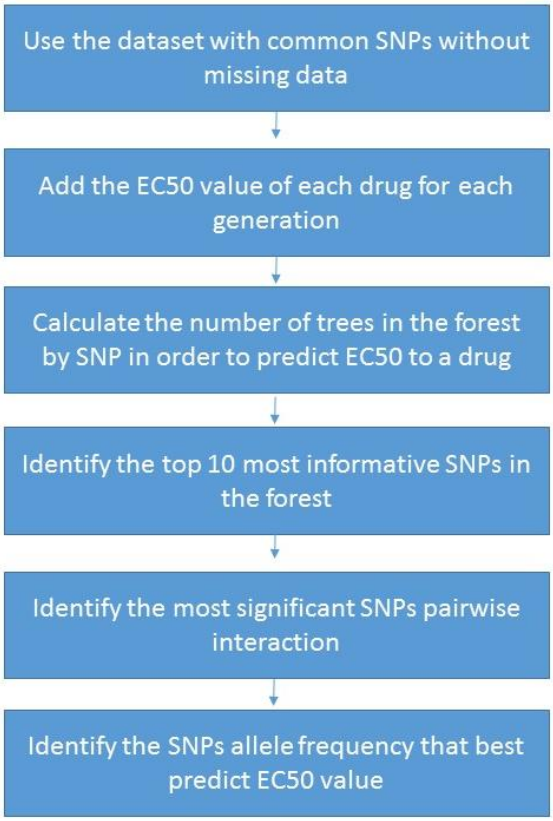

Fig 7. Random Forest prediction analysis pipeline used to predict drug resistance.

4 Results

4.1 SNPs chip construction

4.1.1 SNPs from Barber's Pole Worm

We have retrieved the highest confidence SNPs from the alignments of our DNA sequencing data from two populations (CSIRO IP contributed toward the project) against the two *H. contortus* genome assemblies available at the onset of this study, identifying a repertoire of 13,397,305 SNPs in this species. The 334 selected SNPs forming the initial SNP chip therefore represent 0.0025% of the SNPs potentially available from Barber's pole worm. Clearly, there are an immense number of additional SNP which could be used in future work.

The Illumina sequence data aligned with a read depth of 20X over 44% of the ISE *H. contortus* reference genome. The read trimming and filtering eliminated an average of 5% of reads of which an average of 88% were aligned to the reference genome (Table 4, see Appendix 9.5 for individual NGS libraries statistics). The pre-process filtering removed 66% and 60% of reads in the ISE aligned samples, while 75% of reads aligned to MCM genome were removed. The mean coverage was always above 20 reads per loci. The alignment to the ISE genome (Laing et al., 2012) showed a higher quality alignment compared to alignment with the McM genome (Schwarz et al., 2012) with a lower percentage of reads being rejected during filtering and a much higher coverage (Table 4). This is congruent with the lower quality of the McM genome assembly.

Table 4. Read and contig filtering statistics for the pooled McM and Wal NGS libraries of *H. contortus* aligned to the ISE and MCM reference genomes. See Appendix 9.5 for statistics by individual libraries.

	Reads from Illumina	Remaining after Trimming %	Remaining after alignment to ref genome %	Remaining after filtering reduction %	Remaining after duplication removal %	Remaining after coverage screen %	Mean coverage after screening steps (reads)
TO ISE							
ALL MCM	321051718	94	87	60	94	80	36.02
ALL WAL	286764240	95	88	56	95	82	34.2
TO MCM							
ALL MCM	321051718	94	88	69	94	73	22.92
ALL WAL	286764240	95	89	70	94	72	21.23

The GATK variant calling pipeline identified over 15 million variants of which 84.7% were SNPs (Table 5). The hard filtering step eliminated 12% of SNPs leaving a pool of 11 million high quality SNPs. We did not select SNPs with more than two alleles (multi-allelic SNPs) for the panel.

Table 5. SNP discovery statistics for the combined McMaster1931 and Wallangra2003 samples aligned to the ISE reference genome.

	MCM+WAL to ISE
GATK Variants	15,804,346
SNPs	13,397,305 (84.7%)
% remaining after hard filtering	11,843,880 (88.4%)
% remaining after removing non-biallelic SNPs	11,638,195 (73.6%)
ts/tv ratio	1.85

The selected 250 putatively neutral SNP showed a wide range of allele frequency differences (AFD) between the Wallangra2003 and McMaster1931 populations, with a majority of SNPs showing differences in frequencies between ±0.0 to 0.1, with the extremes exceeding ±0.8 in frequency difference (Fig 8A and 8B). The SNPs also showed an even distribution in minor allele frequencies in both populations (Fig 8C and 8D).

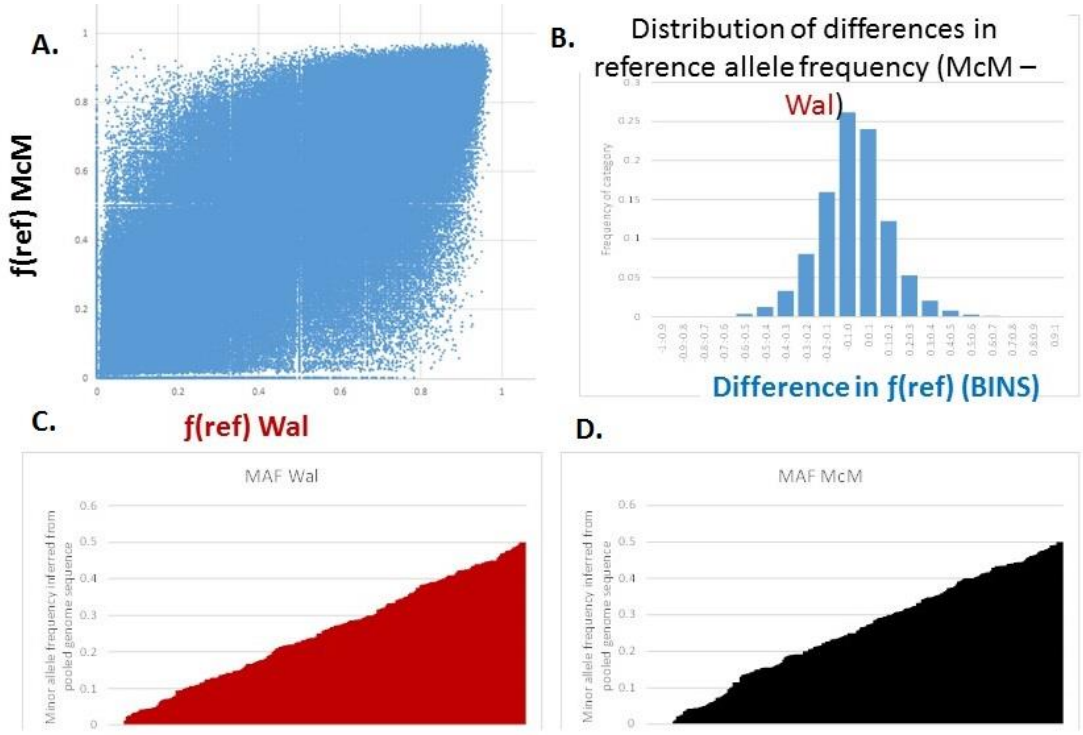

Figure 8. A. Reference allele frequency comparison of selected SNPs between McMaster1931 and Wallangra2003, B. Reference allele frequency differences between McMaster1931 and Wallangra2003 as a frequency distribution histogram (data binned at 0.1 AFD), C. minor allele frequency (MAF) distribution of the selected SNPs in the Wallangra2003 population, D. minor allele frequency (MAF) distribution of the selected SNPs in the McMaster1931 population.

During the project a new *H. contortus* genome assembly was produced. For this report we have mapped the 334 selected SNP to this new genome assembly to refine our knowledge of the position and potential functionality of these SNPs. Of the putatively neutral 250 SNPs selected, 47.5% are located in exons while 44.3% are in non-coding regions (introns or intergenic regions) (Table 6). An additional 8% did not map to the 2018 assembly and therefore do not have an annotation. Of the 88 drug resistance candidate SNPs selected, 94% were found in coding regions of similar resistance genes as previously identified in the ISE assembly (initial and 2018 annotation in SNP table in Appendix 9.2).

Table 6. Annotation of different categories of selected SNPs from the most recent *H. contortus* genome assembly (2018-05-WormBase version).

	Function by location (gene/exon/mRNA) in new genome	**Intergenic or intronic in new genome**	**Other (repeat, low complexity, etc.)**	**Not mapped**
putatively neutral (total= 246)	117	105	4	20
candidate (total= 88)	83	5	0	0
overall (total= 334)	200	110	4	20

The difference in allele frequency between two populations (Wallangra2003 and McMaster1931) was predicted from our NGS data and compared to the results of allelotyping using the sequenom technology for 200 SNPs. For more than 100 of the SNP, the difference between NGS and sequenom was less than 0.05 (Fig 9). This shows that the prediction of allele frequency differences based on sequencing was usually an accurate reflection of the estimates using a more accepted technology. This gave us confidence that our approach of selecting SNP based on allele frequency differences was valid.

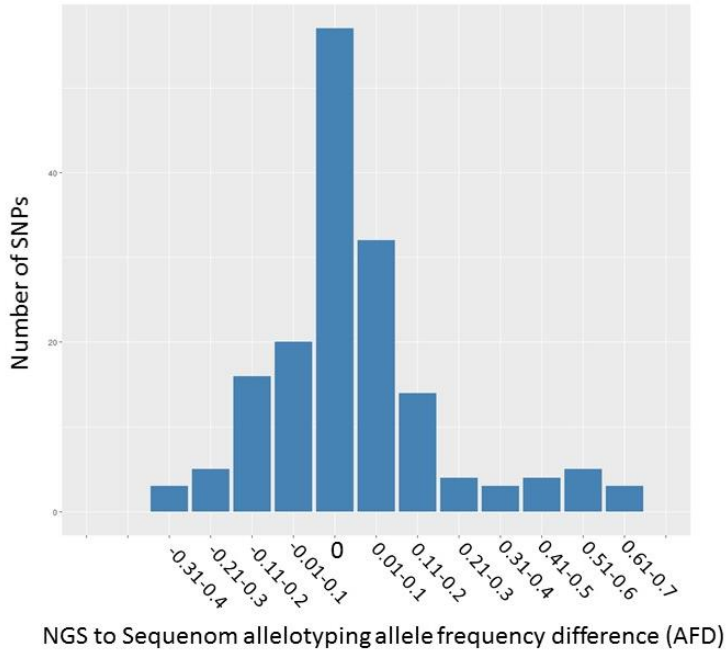

Fig 9. Difference in allele frequency estimation (AFD, frequency difference between Wallangra2003 and McMaster1931 reference allele frequencies) between NGS and Sequenom in 0.1 frequency bins. On the x axis; 0.1 frequency bins, on the y axis; number of SNPs with a specific allele frequency difference.

4.1.2 SNPs chip quality-control filtering

The quality control steps eliminated 22% of the 250 selected SNPs evaluated (Table 7). Some SNP were eliminated for more than one criterion. After the quality control filtering, a total of 197 SNPs remained in the SNP chip. The identity of the SNPs which passed quality control, those which were eliminated and those incompletely evaluated are indicated in Appendix 9.2.

When non-target DNA was amplified by the Sequenom primers more than once from three replicates, the SNP was eliminated from consideration. In table 7, 5% of SNP are shown to have failed this quality control step.

Over all SNPs, there were 9 significant differences (p value <0.05) between the observed and expected frequency at the 50:50 admixture (Table 7 – "admixed population detection"). This low value indicates that the majority of SNP can be evaluated for allele frequency as pools equally as well as for individuals. This is an essential outcome for producing an affordable DNA-test for producers to use. The data obtained from this quality control step are interesting in themselves. The genetically distant populations McMaster1931 and Wallangra2003 showed the highest difference with 0.16 (Table 9) while populations GoldCoast2004 and Goondiwindi2011 showed the lowest difference and

the spatially distant populations Mackay2009 and Cannawigara2006 showing a slightly higher 50:50 difference. Detection of 100%, 75%, 50%, 25% and 0% for the three population pair was also assessed by eye (Fig 10) for each SNP.

Table 7. Quality control filtering statistics from the 334 selected SNPs.

QC step	total number of SNP eliminated	% SNPs eliminated
Multiplex	14	4%
amplification	8	2%
Pooling vs individual	8	2%
Admixed population detection	9	2.5%
co-occuring species	16	5%
Passed	197	78%
Evaluation not complete	84	25%

Table 8. SNPs eliminated due to cross amplification of non-*H. contortus* samples.

Species	Number of SNP amplified
T. circumcincta	16
T. colubriformis	16
O. columbianum	16
O. venulosum	14
O. aries	13
C. ovina	16
C. oncophora	12

From the 250 neutral SNPs, 8 SNPs were eliminated because they showed a higher than 0.15 significant difference allele frequency between pooled and individual samples from either the McMaster1931 or Wallangra2003 populations (Table 7 – "pooling vs individual"). Figure 11 shows a subset of the outcomes from this quality control step. SNPs that showed a non-significant difference of 0.15 in only one of the 2 populations were not eliminated. Most of the SNPs showed that pooled worm samples underestimated the minor allele frequency compared to individual worm samples, though these differences were not statistically significant in most cases.

Table 9. Mean between observed and expected frequency in 50:50 admixture for the 3 population pairs for the 250 putatively neutral SNPs.

Population admixed pair	Mean 50:50 difference	Standard deviation on Mean
Wallangra2003 – McMaster1931	0.164	0.120
GoldCoast2004 – Goondiwindi2011	0.051	0.060
Mackay2009 – Cannawigara2006	0.0645	0.069

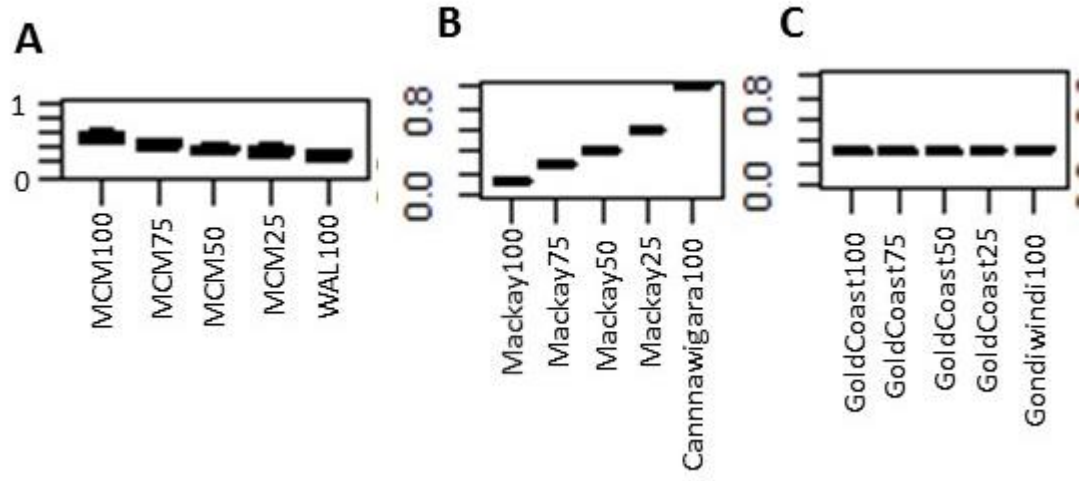

Fig 10. Admixed population detection example (SNP Q_P171221) - ALT allele frequency is shown on the y-axis for all three graphs. (A) Wallangra2003-McMaster1931 pair where the percentage of McMaster1931 (MCM) DNA is indicated on the x-axis. (B) Mackay2009-Cannawigara2006 pair where the percentage of Mackay DNA is indicated on the x-axis. (C) GoldCoast2004-Goondiwindi2011 pair where the percentage of of GoldCoast DNA is indicated on the x-axis. For (A) and (B) the ALT frequency differs between the two isolates and there is an approximately linear relationship between the admixed sample ALT allele frequency and the proportion of the isolate with the highest frequency. In (C) the two isolates have equal ALT allele frequencies, and admixture does not affect the outcome for the admixed samples.

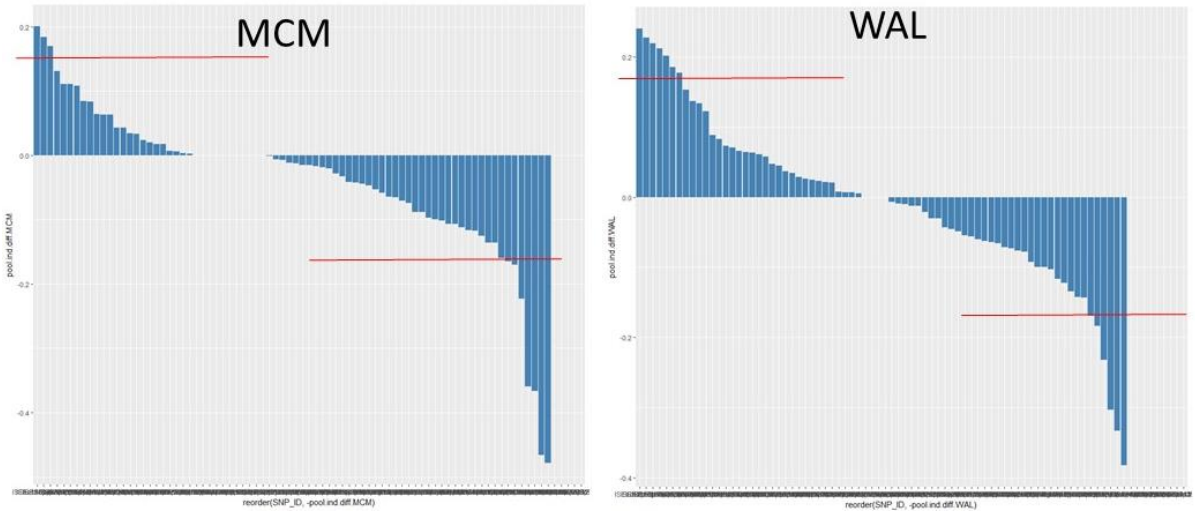

Fig 11. Difference in allele frequency between pooled DNA samples and DNA samples from individuals. Data from the McMaster1931 (left) and Wallangra2003 (right) populations are shown. Red lines indicate the 0.15 difference threshold. SNPs in which the pooled sample overestimated the allele frequency of the individual are represented by blue bars above the axis, SNPs in which the pooled samples underestimated the individual allele frequency are represented with blue bars below the axis. Data shown are from a randomly selected subset of 100 SNP for clarity.

4.2 Spatially dispersed population differentiation

The 34 populations were mapped using a principal component analysis and data from 286 SNPs. The two most explanatory principle components explain 24% of the variation (Fig 12). Principal component axis are unit-less, they represent different dimensions of the data pattern, the values seen along the x and y axis represent the raw component scores (Kassambara 2017). In Fig 12, axis x represents the first dimension explaining the highest proportion of variation and axis y represents the second dimension explaining the second highest proportion of the variation. The mainly lab-derived strain ChisAVRS and the 2 South African isolates are distinctly differentiated by this method. The Dawbuts populations are the most similar at the allele frequency level, especially compared to the Invetus field populations that show 2 clusters.

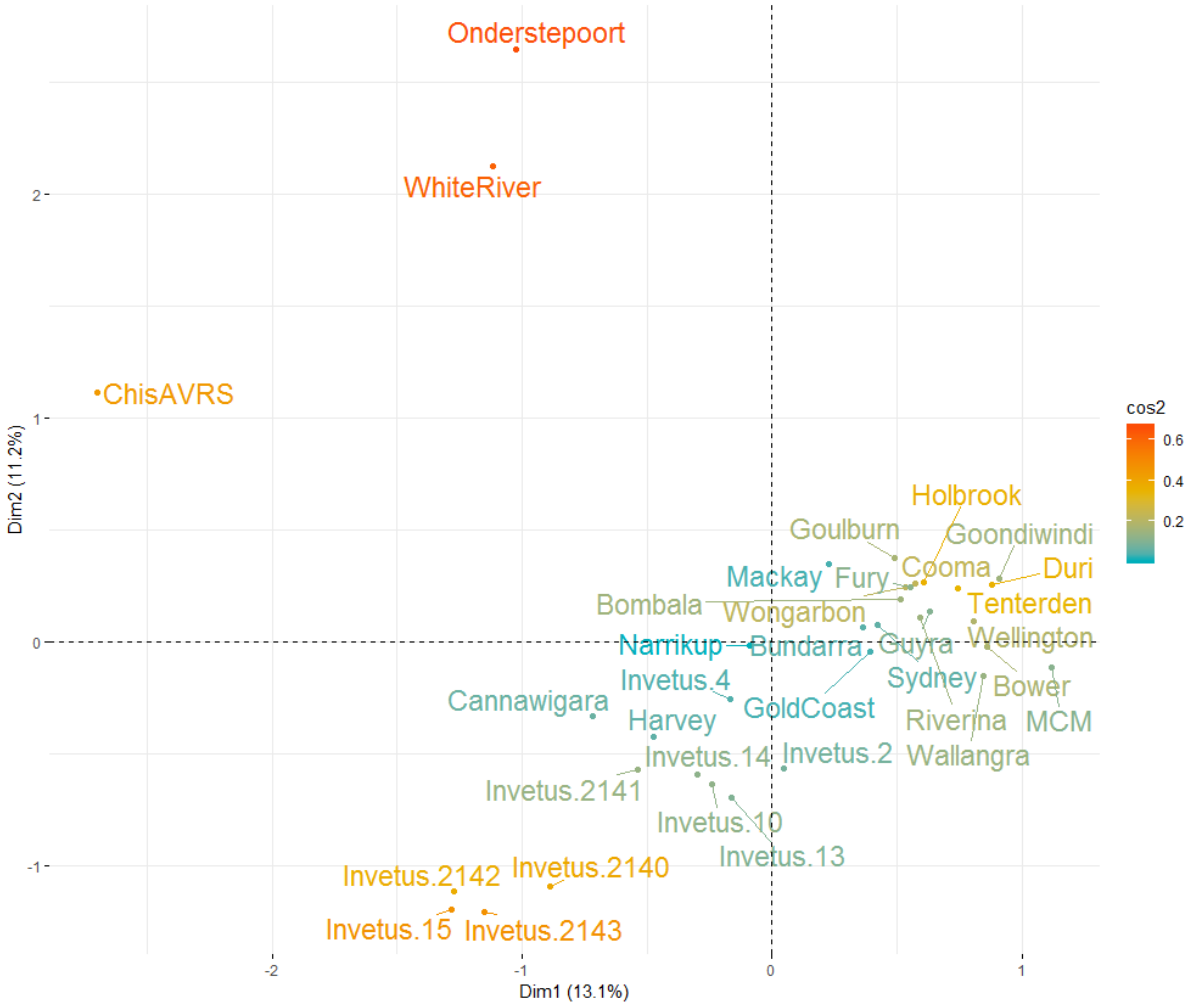

Fig 12. Principal component analysis (PCA) of 34 field populations using 286 SNPs. The x-axis shows the first principle component and the y-axis show the second principle component of the variation observed. The colour shows the cos2 values (see legend) for each data point, and this reflects the distance of the data point from the origin. Cos2 values are indicative of the quality of the placement of the population on the graph, with the greatest quality shown in orange and the poorest in blue. Populations are labelled as in table 2.

The top 10 most informative SNPs are especially useful in discriminating between the 2 largest clusters of field populations (shown in Fig 13 as a hierarchical tree diagram). **The identity of these SNP in combination with the knowledge of what they may be used for is sensitive intellectual property which should be protected.**

Fig 13. Hierarchical clustering of the 34 field populations annotated with the most informative SNPs which differentiate the clusters. Branch lengths (height) are derived from the PCA values scores for each population and cluster (Fig 12). Cluster colouring indicates the PCA optimal grouping number (here 6) to explain the data based on cos2 values (Fig 12). **Figure unavailable in public access version of this report due to IP.**

The PCA was also used to associate an explanatory variable to this pattern. Resistance status and geographical location were used to account for the population patterns on axes x and y respectively. The geographical location could not explain the spatial pattern observed as an ellipse could not be drawn separating geographically distinct populations and populations from very diverse geographical locations were clustered together (Fig 14). However, the genetic data is a better explained by resistance status than by geographical position. Note that the spatial pattern is not necessarily a relationship to geographical location, it refers only to spatial patterns in the data and in this case is better explained by drug resistance status than geographical origin. There is an overlap between resistant and susceptible, however the ellipse of susceptible populations is much narrower indicating a precise and narrow distribution.

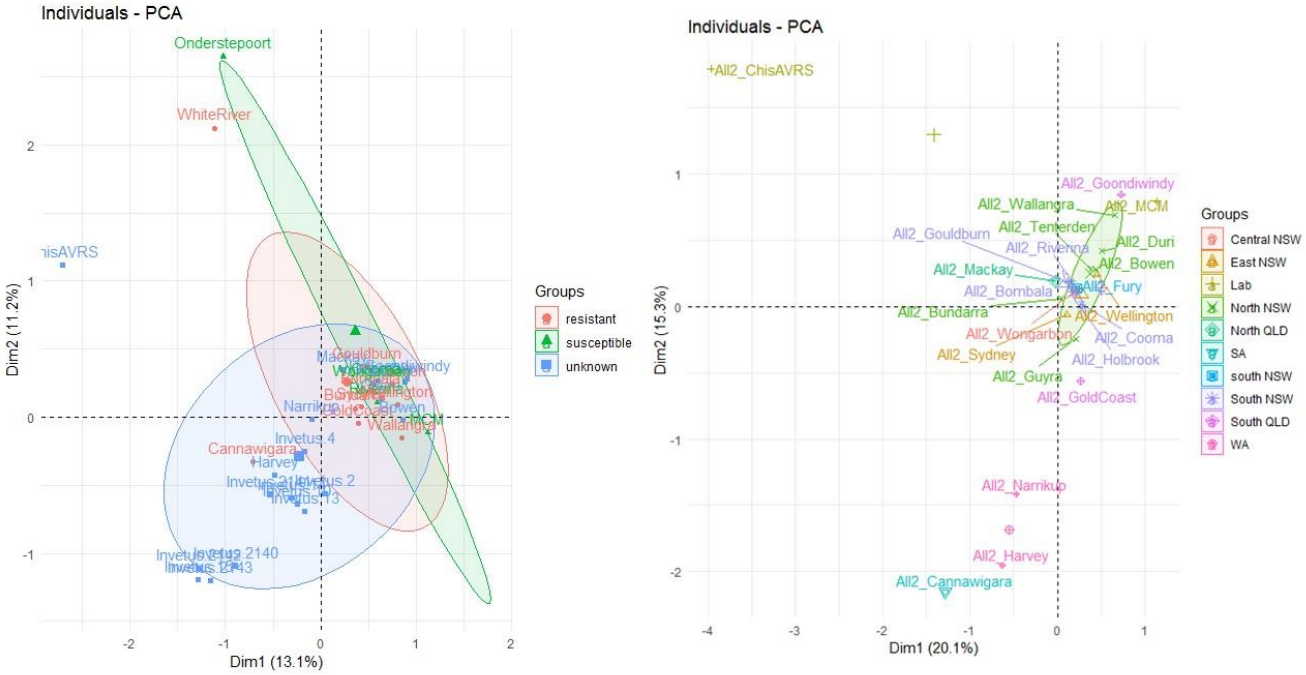

Fig 14. Association of PCA pattern with explanatory variables. Left, using resistant, unknown and susceptible categories. Right, using the geographical location as explanatory variable with 22 populations with known Australian geographical location. The X axis represents the first dimension of the PCA representing the highest variation possible with axis score, while Y axis represents the second dimension of the PCA explaining the second highest percentage of variation in the data with axis score. Note that the figure on the left shows an interpretable association between resistance/susceptibility categories and PCA pattern whilst the figure on the right shows that no geographical category is associated with the PCA pattern.

The allele frequencies from all shared SNPs between populations were used to compare the minor allele frequency diversity between populations using the Shannon H index (Fig 15). The Shannon index shows significant differences between multiple populations; the South African and ChisAVRS populations have a significantly lower allele frequency diversity (meaning MAF closer to 0 for most SNPs) compared to field-derived Australian populations. Within the field populations, we observe

significant differences between pairs of populations exhibiting moderate diversity (e.g. Goondiwindi2011 and GoldCoast2004) and those exhibiting high diversity such as McMaster1931 and Wallangra2003. The minor allele frequencies of each SNP can be interpolated into genotypes, which can then be mapped into a network of shared genotypes between populations (Fig 15). This network shows that genotype MLG.9 is shared among multiple populations while the genotypes found in the 2 South African populations (in dark blue) are mostly not shared with other Australian populations. The usefullness of the genotype network will be realized when sufficient field populations with associated resistance status are compiled, detecting shared genotypes which distinguish shared drug resistance.

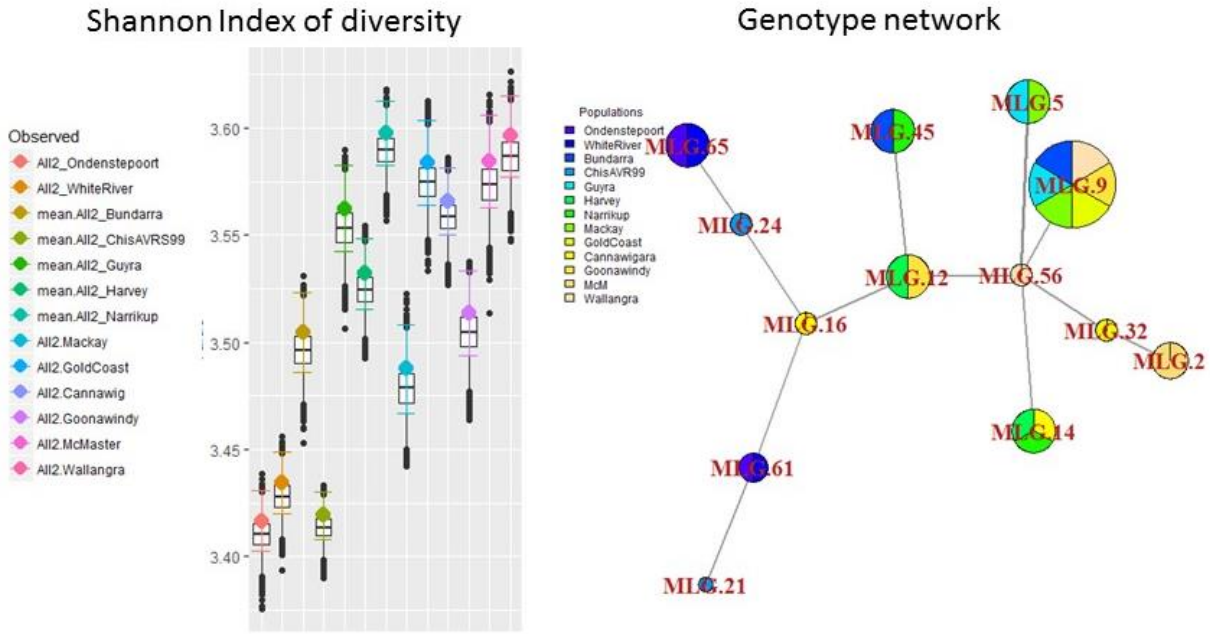

Fig 15. Left, Shannon diversity index used to compare allele frequencies between populations of *H. contortus*. Right, multi-locus genotype network of field populations showing the proportion of each *H. contortus* population contributing to the genotype as differing colours in the pie graphs. The lines between genotypes indicate the genetic distance between them, so that pairs which are closely related (e.g. MLG.2 and MLG.32) are closer together on the diagram than those which are more distantly related (e.g. MLG.65 and MLG.2).

The overall heterozygosity was lowest in the lab-derived ChisAVRS population with 0.13 and highest in the field-collected population Wongarbon with 0.30 (Fig 16); Wongarbon is one of the isolates contributed by Dawbuts. The McMaster1931 population (McM in Fig16), has a heterozygosity value as high as some of the Dawbuts field populations, which is unexpected given its history, but in line with genotyping undertaken in the past with different methodology (Hunt *et al* 2008). Given the large 95% confidence intrervals for each populations (on average 0.16), there were no significant differences based on heterozygosity between the spatial dataset populations.

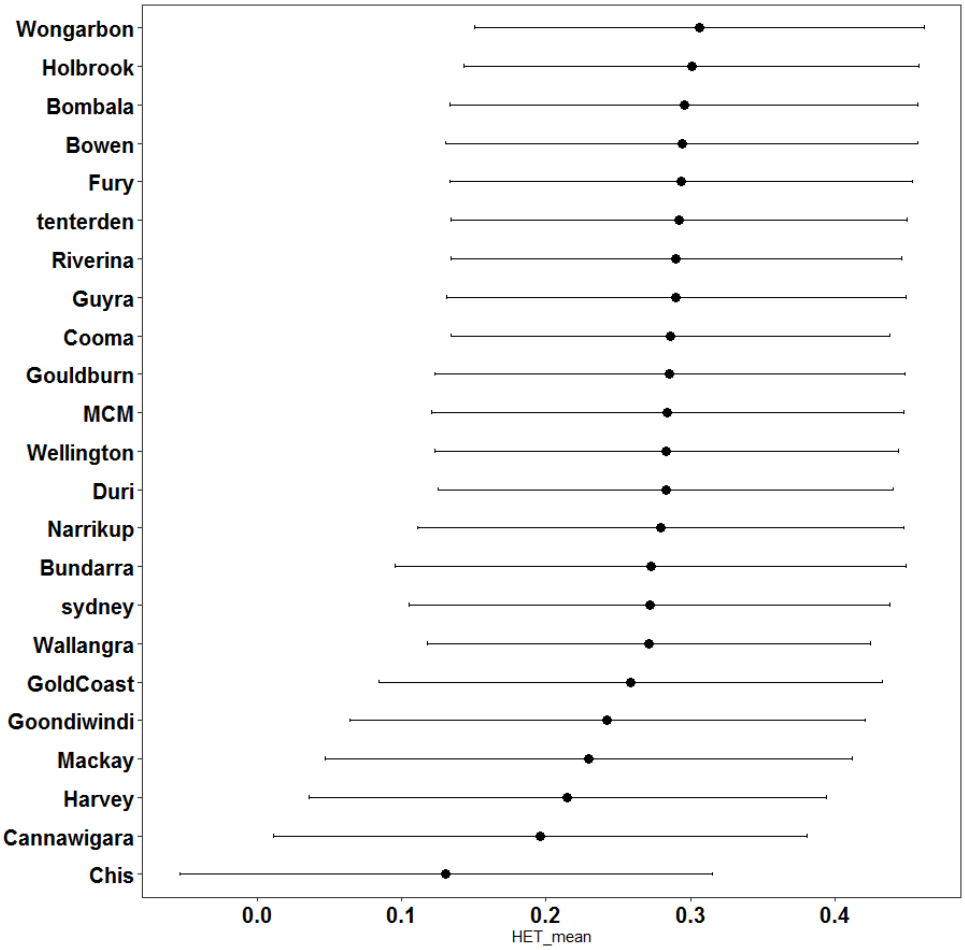

Fig 16. Population heterozygosity comparison with 95% confidence intervals based on minor allele frequency from 68 SNPs.

4.3 *In vivo* drug selection experiment

4.3.1 Increase in resistance with drug selection

To demonstrate the use of SNPs to detect selection for drench resistance within sheep enterprises, an experiment was conducted where susceptible and resistant BPW were mixed and used to infect sheep. The BPW populations in these sheep were transferred between sheep each generation (passaging), and each generation was named G1 through to G6, whilst the parental isolates were designated G0. In generations G3 to G6 drenches were used to select resistant survivors (SEL). Control populations were also passaged with no drug selection in G3 through to G6 (UNS). Two experiments were conducted using different resistant isolates mixed with the McMaster1931 susceptible isolate. The first experiment (BxM) used the multiple drug resistant isolate Bundarra2017, and drench selection using only monepantel. The second experiment (WxM) used the multiple drug resistant isolate Wallangra2003, and selection with a combination of oxfendazole, abamectin and closantel. Over time the response to selection was assayed in two different ways, first using the LDA test (for abamectin, monepantel, thiabendazole and levamisole) and second using the difference in WEC before and after drenching.

4.3.1.1 *Single drug selection (BxM experiment)*

In the BxM experiment, the Bundarra2017 isolate was shown to be more resistant to all four drugs assayed in the LDA, compared to McMaster1931: Bundarra2017 was 629, 6.45, 700 and 58.1 fold more resistant than McMaster1931 for thiabendazole, abamectin, monepantel and levamisole, respectively. The dose-response in LDAs for monepantel is complex, and in some cases best interpreted as two sub-populations with differing EC50 values (Kotze et al., 2018; Fig 17). After mixing Bundarra2017 and McMaster1931, the G2 generation had dose response curves with two populations evident: one with 1.39 fold resistance, statistically not different from McMaster1931, and another with 2611 fold resistance over McMaster1931 (mean values for 3 lines passaged in sheep: Fig 17). In the unselected experiment, this biphasic phenomenon persists through till G6 (Fig 18) where the resistance ratios are 1.55 and 1321 fold over McMaster1931 for the two sub-populations (mean values for 3 lines passaged in sheep). In contrast, when monepantel selection is applied, the sub-population with higher susceptibility disappeared between G4 and G5, and at G6 all the larvae appear highly resistant at 1561 fold higher than McMaster1931 (mean values for 3 lines passaged in sheep: Fig 18). LDA results are tabulated in appendices 9.6 (Raw LDA data) and 9.7 (EC50 comparison data).

LDAs were used to follow drench compounds not under selection in the BxM experiment to see if the selection for monepantel resistance was independent from selection for other drench resistances. For levamisole, the resistance ratios relative to McMaster1931 between the SEL and UNS experiments did not differ at G6; 47.9 and 59.9 for G6SEL and G6UNS respectively. Similarly for abamectin, the resistance ratios relative to McMaster1931 between the SEL and UNS experiments did not differ at G6; 3.76 and 3.97 for G6SEL and G6UNS respectively. However for thiabendazole a difference between G6SEL and G6UNS was observed; 6.56 and 1.44 for G6SEL and G6UNS respectively.

For three generations, WEC data were obtained before and after drenching the animals. For the BxM experiment, the apparent mean efficacy (n=3) calculated from these values was 30%, 3% and 0% for generations 3, 4 and 5 respectively. In these calculations, where WEC increased after dosing, efficacy was interpreted as 0%. The low sample size means that this data should be viewed as indicative only, but nevertheless, it shows that the treatment and selection regime was effective at increasing the proportion of resistant individuals in the population, in line with the findings from the LDA tests.

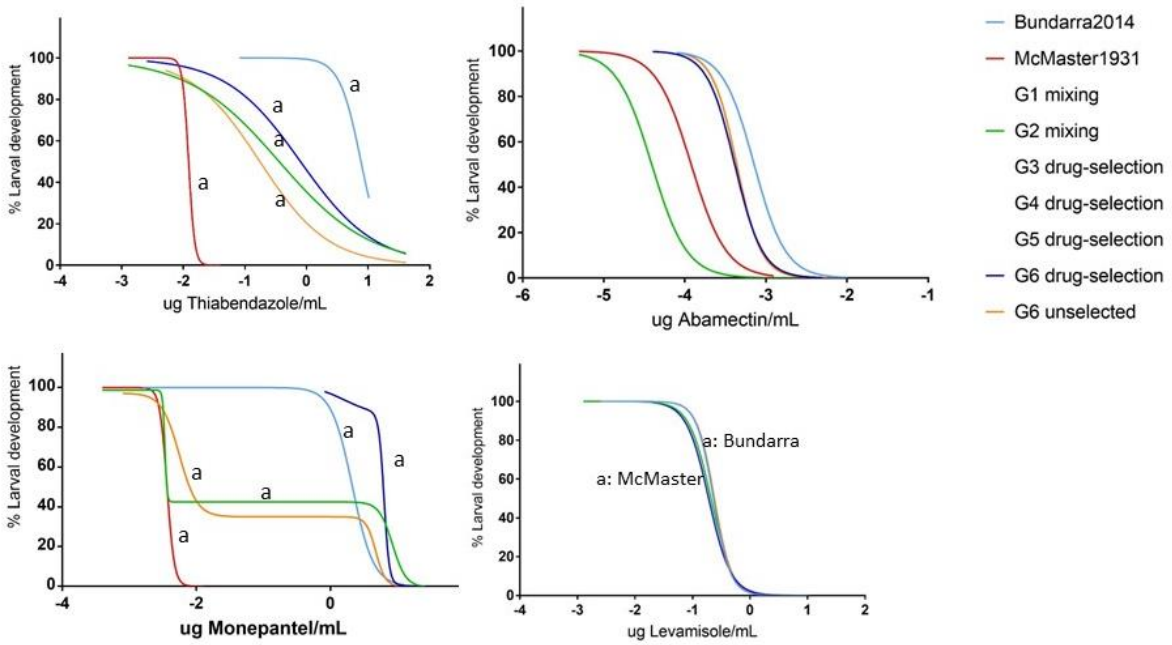

Fig 17. Dose response curves from the larval development assay tested on 4 drugs from the BxM line. Curves with the same letter represent significant differences in EC50. Log10 drug dose is on x-axis while percentage of larvae that developped in the precense of the drug is on y-axis.

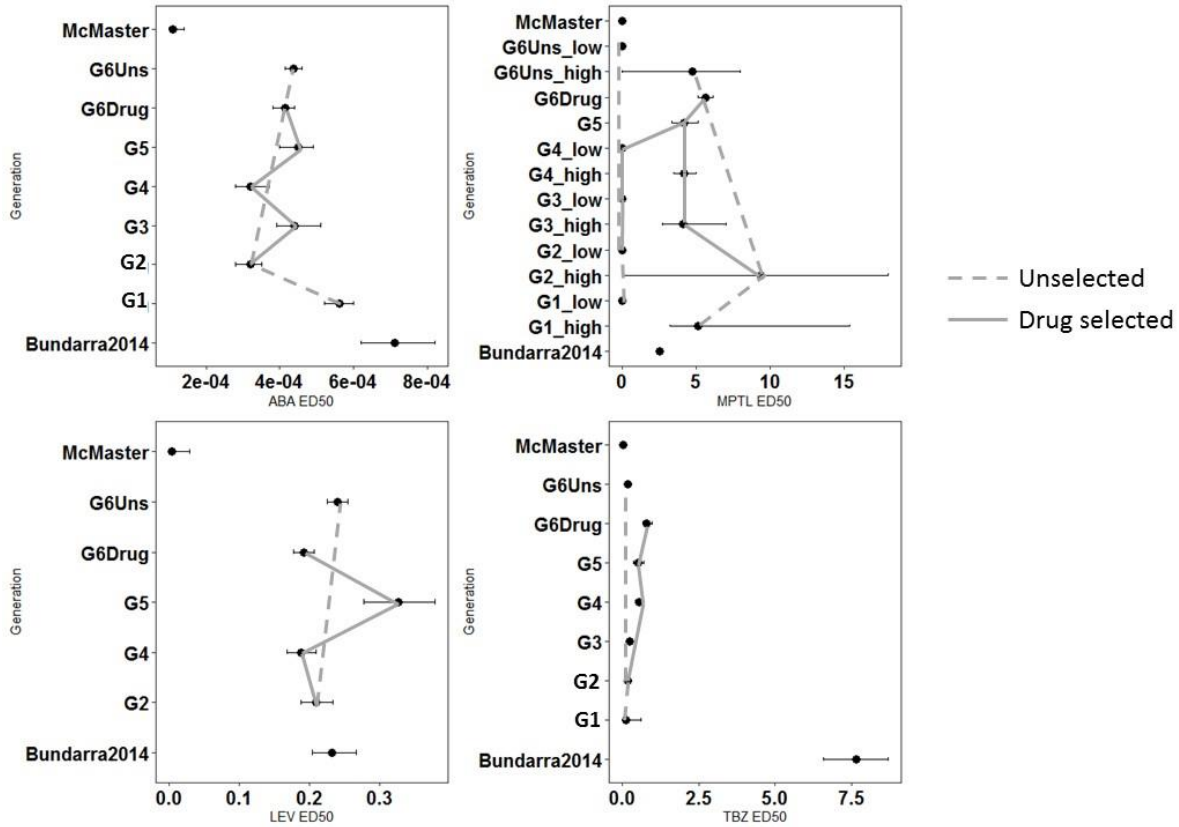

Fig 18. EC50 with 95% confidence intervals on the BxM line for 4 drugs. EC50 points from the same treatment (drug selected vs unselected) connected by full and dotted line respectively as generation number increases on the y axis. Some very narrow 95% confidence intrvals are not visible on the right and left side of points due to the x-axis scale needed to show all populations in the same graph. ABA, abamectin; LEV, levamisole; TBZ, thiabendazole; MPTL, monepantel.

4.3.1.2 Multiple drug selection (WxM)

In the WxM experiment, the Wallangra2003 isolate was used. This isolate had been shown in the past to be multiply resistant to macrocyclic lactones, benzimidazoles, closantel and levamisole (Love et al., 2003; Hunt et al., 2010). We planned to select for resistance to multiple drugs in the WxM experiment, so before the experiment began, two lines of Wallangra2003 were produced using a single passage with drug selection to enrich for resistant individuals. Wallangra2003OXF was generated using an *in vivo* treatment with the product Flukazole (Virbac) containing oxfendazole and the anti-*Fasciola* drug triclabendazole. Wallangra2003ABACLO was generated using an *in vivo* treatment with the product Avomec Duel (Merial) containing closantel and abamectin. The Wallangra2003OXF isolate was shown in LDAs to be more resistant to abamectin and thiabendazole, compared to McMaster1931, but not more resistant to monepantel: Wallangra2003OXF was 6.00 and 4.55 fold more resistant than McMaster1931 for thiabendazole and abamectin, respectively (Fig 19). The Wallangra2003ABACLO isolate was shown in LDAs to be more resistant to abamectin and monepantel, compared to McMaster1931, but not more resistant to thiabendazole: Wallangra2003ABACLO was 7.18 and 2.00 fold more resistant than McMaster1931 for abamectin and monepantel, respectively but the monepantel result is not significant statistically (Fig 19). A mixture of Wallangra2003ABACLO and Wallangra2003OXF was tested in LDA tests and found to be 141 fold more resistant to levamisole than McMaster1931. After mixing Wallangra2003ABACLO, Wallangra2003OXF and McMaster1931 at a ratio of 1:1:2, the G2 generation retained resistance to abamectin and levamisole of 4.64 and 75.2 fold greater than McMaster1931 when analysed by LDA (mean values for 3 lines passaged in sheep). No resistance against thiabendazole was evident in the G2. In the UNS experiment these resistance ratios remained unchanged through to G6, when resistance to abamectin of 4.09 fold greater than McMaster1931 was evident, with no resistance to monepantel or thiabendazole (mean values for 3 lines passaged in sheep Fig 19). In contrast, the SEL experiment using the closantel/oxfendazole/abamectin combination produced G6 populations with mean resistance levels of 3.18 and 6.46 fold higher than McMaster1931 for thiabendazole and abamectin respectively (mean values for 3 lines passaged in sheep Fig 20). It is not possible to assay for closantel resistance in LDA tests. LDA results are tabulated in appendices 9.6 and 9.7.

LDAs were used to follow drench compounds not under selection in the WxM experiment to see if the selection for resistance to the closantel/oxfendazole/abamectin combination was independent from selection for other drench resistances. The change in monepantel sensitivity from G2 to G6 was negligible for both the SEL and UNS lines, showing no relationship between resistance to this compound and the others used in the experiment (Fig 20). For levamisole, the resistance levels were 64.9 fold greater than McMaster1931 at G6 in the UNS lines and 88.2 fold greater than McMaster1931 at G6 in the SEL lines; these values are not statistically different, showing that levamisole sensitivity is also independent of selection for resistance to the other compounds used (Fig 20).

For three generations, WEC data were obtained before and after drenching the animals. For the WxM experiment, the apparent mean efficacy (n=3) calculated from these values was 30%, 33% and 13% for generations 3, 4 and 5 respectively. In these calculations, where WEC increased after dosing,

efficacy was interpreted as 0%. The low sample size means that this data should be viewed as indicative only, but nevertheless, it shows that the treatment and selection regime was effective at increasing the proportion of resistant individuals in the population, in line with the findings from the LDA tests. The selection toward resistance was slower in the WxM experiment compared to the BxM experiment, possibly reflecting the more complex genetics underlying the multiple drug resistance under selection in WxM.

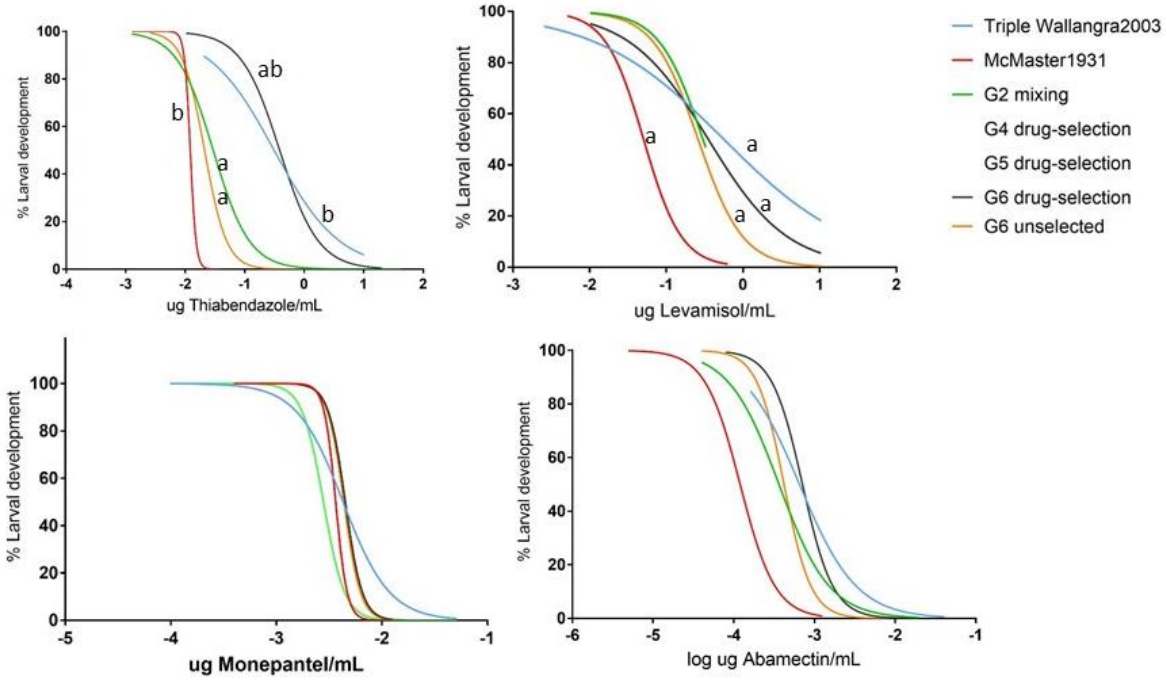

Fig 19. Dose response curves from the larval development assay tested on 4 drugs from the WxM line. Curves with the same letter represent significant differences in EC50. Log10 drug dose is on x-axis while percentage of larvae that developped in the precense of the drug is on y-axis.

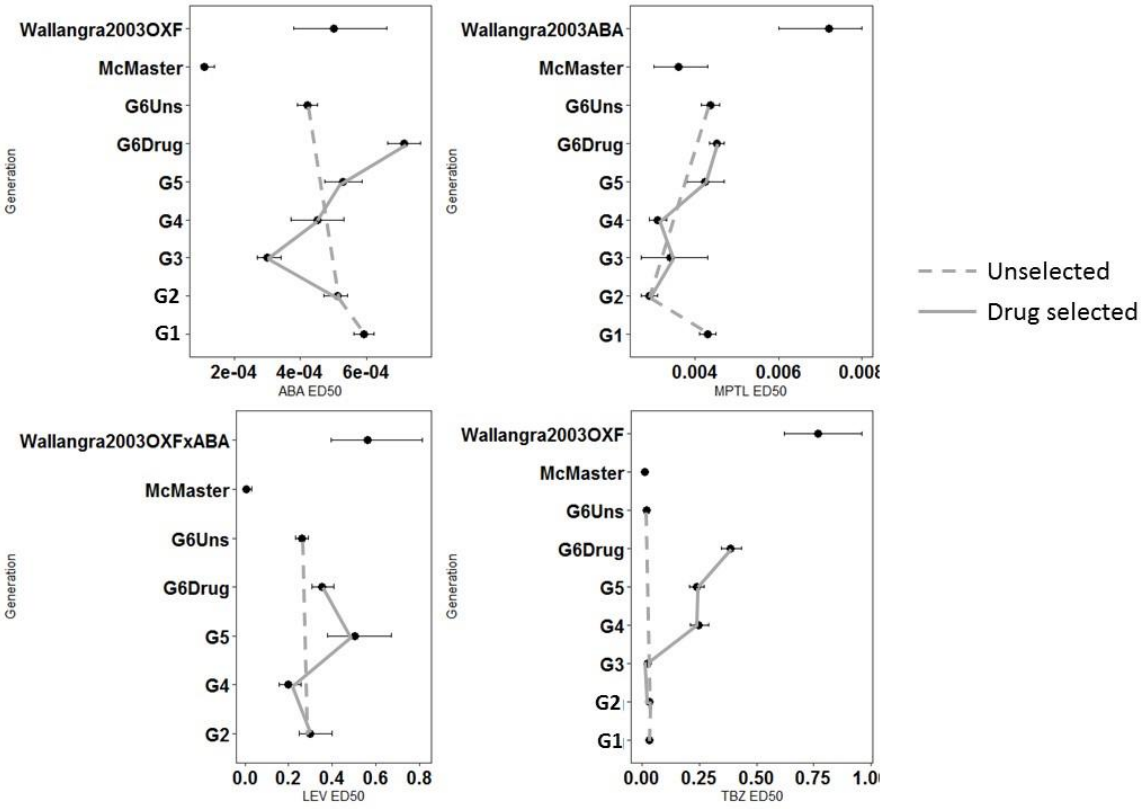

Fig 20. EC50 with 95% confidence intervals on the BxM line for 4 drugs. EC50 points from the same treatment (drug selected vs unselected) connected by full and dotted line respectively as generation number increases on the y axis. ABA, abamectin; LEV, levamisole; TBZ, thiabendazole; MPTL, monepantel.

4.3.2 Association between allelotyping results and drench resistance

Samples of BPW from each passage of the longitudinal experiments BxM and WxM were allelotyped and this data used to associate DNA-test outcomes with drench resistance phenotypes. This analysis was undertaken at the global or whole genome (whole set of SNP assayed) level and also at the individual SNP level. The global level of analysis was undertaken to see if there were any effects of drench selection which had a large effect across the data set. These two levels, global versus individual SNP, will need to be taken into account when the research is translated into predictive tools for sheep producers to use, as the distinction between drench resistance selection and the influx of other populations via migration will need to be established. The individual level analysis can be undertaken using many different statistical approaches. We have used four methods to identify SNPs that may be predictive of drench resistance in the populations studied. These individual SNPs analyses are described in two different sections for the two selection experiments.

4.3.2.1 Global change in BPW genetics during the BxM and WxM experiments

Samples of parasites from the longitudinal experiments BxM and WxM were collected and allelotyped using the Sequenom platform as described (Methods section 3.2). The raw allelotyping results for each dataset and each SNP are available in Appendix 9.8.

The level of heterozygosity of BPW lines in the WxM and BxM experiments changed over time (Fig 21, 22). The SEL lines exhibited a decrease in heterozygosity, markedly between G5 and G6, showing as expected that a small decrease in whole genome heterozygosity accompanies increasing drug resistance. The heterozygosity of the SEL lines however is not significantly different from the heterozygosity of the UNS lines. The changes observed in the mean global heterozygosity (data not shown) values are small in comparison to the total variation between lines within the SEL and UNS groups for both BxM and WxM, showing that the signatures for drug resistance are more specific to fewer markers within these populations. The mean global heterozygosity values showed a 0.05 difference between the UNS and SEL samples in the WxM G6 and G4 compared to G3, however not significant at this point. The heterozygosity of the *H. contortus* populations for the different passages of the BxM experiment are illustrated by pairwise frequency histograms in Fig 20, and for the WxM experiment in Fig 22. In these diagrams the ALT allele frequency for each SNP is combined into bins of 0.05 of size so that 20 bins are illustrated in each histogram. The bars toward the left and right extremes of the histograms show the number of SNPs with lower levels of heterozygosity, and bars toward the centre have the highest degree of heterozygosity. These diagrams illustrate the observations for over-all heterozygosity change with selection, especially for the WxM experiment. Another interesting observation is the decrease in heterozygosity between G1 and G2 which occurs in the BxM experiment.

Fig 21. Allele frequency histograms using 0.05 bins for the BxM dataset. Each graph represents an overlay of 2 graphs with separate bar colour (blue or bright green), when 2 bars overlap, the resulting colour is dark green. From left to right and upper to lower the histograms compare (1) McMaster1931 (MCM, blue) with Bundarra2017 (BUN, green), (2) the G1 (blue) and G2 (green) generations of the initial mixing phase of the BxM experiment, (3) the G3 SEL (blue) with G3 UNS (green), (4) the G4 SEL (blue) with G4 UNS (green) and (5) the G5 SEL (blue) with G5 UNS (green).

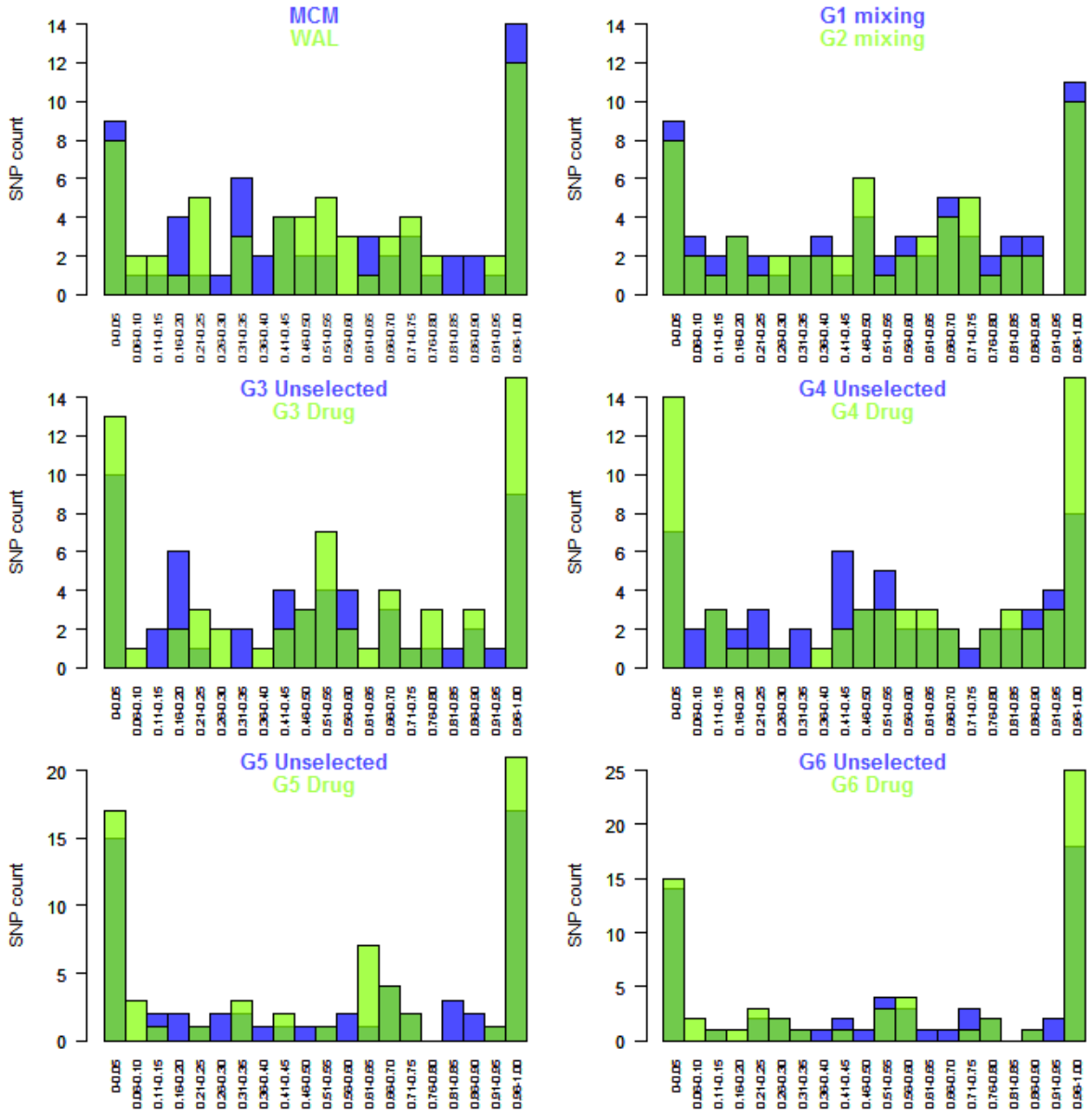

Fig 22. Allele frequency spectrum for the WxM data. Each graph represents an overlay of 2 graphs with separate bar colour (blue or bright green), when 2 bars overlap, the resulting colour is dark green. For the G3, G4, G5 and G6 graphs, the blue bard represent the unselected samples while the bright green bars represent the drug-selected samples. Note the maximum value of the y-axis is greater for G5 and G6 graphs than the others.

A second way of comparing the generations of the selection experiment globally across all SNPs is to calculate a statistic called Fst which can be used to compare pairs of samples. Tables 10 and 11 show the Fst values for pairs of samples from the WxM and BxM experiments, including p-values for significant differences. The divergence between the parental isolates is significant in both experiments, and the divergence between the SEL and UNS lines can be followed over multiple passages. For the WxM experiment the SEL vs UNS divergence increases as selection for drug resistance progresses, however this is less clear for BxM. Similarly the divergence of the SEL lines from the WxM experiment from the McMaster1931 parent samples becomes progressively larger and p-values decrease as selection for the drug resistance proceeds. The UNS lines also diverge from the McMaster1931 parent, however this is markedly less distinct compared to the SEL lines. Similarly, the SEL and UNS lines in the WxM experiment diverge from the Wallangra2003 parent, but the magnitude of the Fst values is less at G6. The BxM experiment shows a similar pattern but with lower Fst values and more gradual change over passages (Table 11). Note that G5 data is missing from the BxM experiment due to a technical issue with the allelotyping.

Table 10. Pairwise Fst statistics for the WxM multiple drug selection experiment

Comparison	Fst value	p-value
WAL[1] vs MCM[2]	0.177	0.000
G3 SEL vs G3 UNS	0.006	0.414
G4 SEL vs G4 UNS	0.046	0.040
G5 SEL vs G5 UNS	0.238	0.000
G6 SEL vs G6 UNS	0.250	0.000
MCM vs G1	0.153	0.000
MCM vs G2	0.153	0.000
MCM vs G3 UNS	0.131	0.000
MCM vs G4 UNS	0.111	0.000
MCM vs G5 UNS	0.311	0.000
MCM vs G6 UNS	0.233	0.000
MCM vs G3 SEL	0.228	0.000
MCM vs G4 SEL	0.301	0.000
MCM vs G5 SEL	0.372	0.000
MCM vs G6 SEL	0.456	0.000
WAL vs G1	0.031	0.128
WAL vs G2	0.046	0.068
WAL vs G3 UNS	0.032	0.122
WAL vs G4 UNS	0.031	0.138
WAL vs G5 UNS	0.167	0.000
WAL vs G6 UNS	0.103	0.002
WAL vs G3 SEL	0.041	0.06
WAL vs G4 SEL	0.080	0.004
WAL vs G5 SEL	0.143	0.000
WAL vs G6 SEL	0.245	0.000

1 – Wallangra2003 *H. contortus* isolate, 2 – McMaster1931 *H. contortus* isolate

Table 11. Pairwise Fst statistics for the BxM multiple drug selection experiment

Comparison	Fst value[3]	p-value
BUN[1] vs MCM[2]	0.158	0.000
G3 SEL vs G3 UNS	0.000	1.000
G4 SEL vs G4 UNS	0.000	0.996
G6 SEL vs G6 UNS	0.026	0.108
MCM vs G1	0.042	0.012
MCM vs G2	0.223	0.000
MCM vs G3 UNS	0.053	0.006
MCM vs G4 UNS	0.052	0.004
MCM vs G6 UNS	0.132	0.000
MCM vs G3 SEL	0.063	0.002
MCM vs G4 SEL	0.095	0.000
MCM vs G6 SEL	0.187	0.000
BUN vs G1	0.019	0.192
BUN vs G2	0.168	0.000
BUN vs G3 UNS	0.012	0.270
BUN vs G4 UNS	0.011	0.256
BUN vs G6 UNS	0.105	0.000
BUN vs G3 SEL	0.007	0.270

1 – Bundarra2017 *H. contortus* isolate, 2 – McMaster1931 *H. contortus* isolate, 3 – Fst values which were negative are shown as zero.

A third way of comparing the generations of the selection experiment globally across all SNPs is to use the data in a principal component analysis (PCA) to see if the SEL and UNS lines are clearly differentiated from one another and the G1 and G2 MIX lines in the data space. Further details of the PCA method can be found in the methods section. For the BxM experiment, the PCA does not show a clear SEL vs UNS cluster but the mapping of the lines as an explanatory factor does indicate differentiation on an axis additional to the first 2 PCA axes (Fig 23). This could indicate that the *in vivo* selection for monepantel resistance in the BxM experiment did not continue for enough passages to see a clear differentiation across the whole genome using component analysis. In contrast, the WxM experiment PCA is more completely resolved and ellipses separated drench resistant and drench susceptible samples (Fig 24). This is another indication that multiple drug selection can have global effects on the genome whereas single compound selection, such as that conducted in the BxM experiment has a smaller effect on the whole genome.

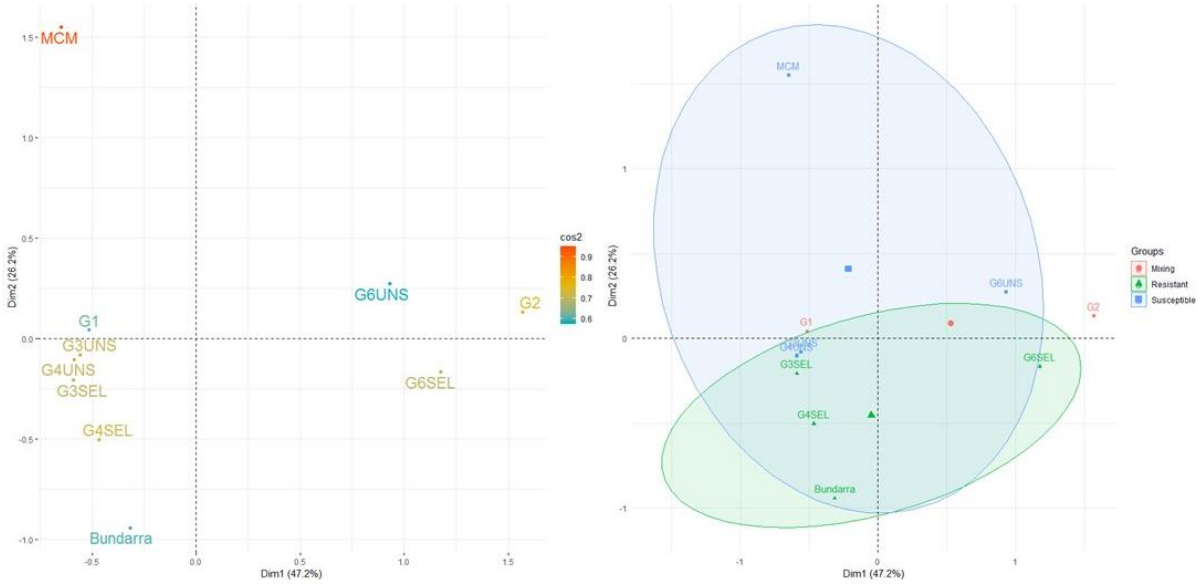

Fig 23. Principal component analysis (PCA) of population distribution for the BxM lines. On the left, the different samples are plotted against the first two principle components, Dim1 and Dim2 (these have no units), the proportion of variation explained by the components is indicated on the axes labels. On the right, the output of automated analysis which draws ellipses around data groups evaluate if there is a genetic segregation between the defined groups, in this case the drench resistant (SEL samples and resistant parent isolate), the drench susceptible (UNS samples and the susceptible parent isolate) and the "mixed" samples from G1 and G2. In this example there is a tendency toward separation between SEL and UNS, but this is not statistically supported due to the extensive overlap with the "mix" group of samples from generations 1 and 2.

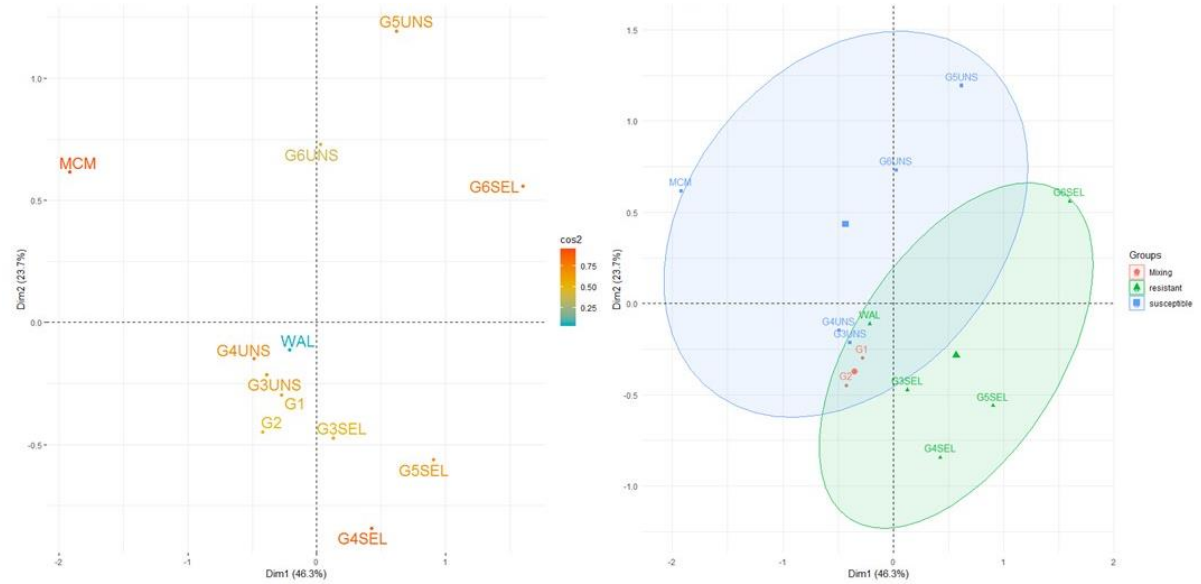

Fig 24. Principal component analysis (PCA) of population distribution for the WxM lines. On the left, the different samples are plotted against the first two principle components, Dim1 and Dim2 (these have no units), the proportion of variation explained by the components is indicated on the axes labels. On the right, the output of automated analysis which draws ellipses around data groups evaluate if there is a genetic segregation between the defined groups. The mixed G1 and G2 passages are shown in red text, whilst the drug selected SEL lines with the resistant parent isolate (labelled "resistant") are encapsulated by the green ellipse and the unselected UNS lines with the susceptible parent isolate (labelled "susceptible") are encapsulated by the blue ellipse.

4.3.2.2 SNPs markers of resistance (BxM experiment)

The identity of these SNP from the four analyses described here, in combination with the knowledge of what they may be used for is sensitive intellectual property which should be protected.

To identify the SNPs most contributing to the observed changes in resistance to monepantel over time in the BxM experiment, four analyses were conducted.

Firstly an analysis (ANOVA method) was performed in which G1 and G2 were identified as "mix", all the drug selected samples identified as "drug" and all the unselected samples identified as "uns". ANOVA followed by Tukey test was performed to identify the SNPs that showed significant differences in the allele frequency between "drug" and "uns". Four SNPs were identified (Fig 25) with significant differences ($p<0.05$).

Fig 25. SNPs significantly different between monepantel resistant and susceptible based on allele frequency. **The identity of these SNP, in combination with the knowledge of what they may be used for is sensitive intellectual property which should be protected. Figure unavailable in public access version of this report due to IP.**

In the second method (FST method) a statistic (Fst) was calculated for each SNP in each generation of the BxM experiment, in this case the Fst values are not pairwise comparisons between samples, but rather are comparisons between an individual SNP and the other SNPs assayed within the same sample. Using this approach a "signature of selection" can be observed using data from the BxM experiment with one SNP identified as possibly contributing to monepantel resistance based on the Fst value at generation 6 in the SEL lines compared to the G6 UNS lines.

The third approach for identifying SNPs associated with drench resistance selection in the BxM experiment depended on the PCA strategy (PCA method - Fig 23). The SNPs identified contributing most to the difference between drench resistance and susceptibility were identified (Table 12).

Table 12. SNP associated with selection with monepantel from analysis of the BxM experiment.

Table unavailable in public access version of this report due to IP.

A fourth method for associating SNP with drench resistance used the random forest statistical methodology (RF method). An advantage of this method is that the magnitude of resistance (from the LDA data) can be used, rather than treating resistance in the categorical way used for other methods. A disadvantage of the method is that it cannot handle missing data, so only 95 of the SNP could be utilised in the analysis. Monepantel EC50 values between 0.5 µg/mL and 1.5 µg/mL could be predicted using the SNP in the BxM experiment (Fig. 26A).

The RF method can also be used to find associations with categorical data, and this analysis was also undertaken, comparing the samples from drench selected populations (SEL) to those from unselected populations (UNS) (Fig 26B).

Fig 26. Heat Map from the BxM experiment for the RF method. **Figure unavailable in public access version of this report due to IP.**

In summary from the BxM experiment, 14 SNPs are associated with monepantel selection, and five of these were significant using two separate statistical methodologies. No SNP were identified in more than 2 of the statistical methodologies. Although some of these SNPs could have a causative relationship with monepantel resistance, it is more appropriate to consider these simply as associated with resistance to monepantel in the Bundarra2017 population, as some are likely associated with resistance for reasons other than causation, such as genomic linkage with causative SNP or with genetic modifiers that enhance survival or resistant genotypes.

Because the RF methodology is able to incorporate the EC50 data from the LDAs, we were able to use the data from the BxM experiment to look for associations between SNP and the other anthelmintics not used for selection. This was possible as there was variation observed between samples in their survival response to thiabendazole, levamisole and abamectin (Fig 18). Significant associations were detected between SNPs and EC50 for levamisole (3 SNP) and abamectin (2 SNP), but not for thiabendazole (Fig 27).

Fig 27. **Figure unavailable in public access version of this report due to IP.**

4.3.2.3 SNPs markers of resistance (WxM experiment)

The identity of these SNP from the four analyses described here (except the BTUB SNP), in combination with the knowledge of what they may be used for is sensitive intellectual property which should be protected.

To identify the SNP contributing most to the observed changes in resistance to the combined treatment with oxfendazole, abamectin and closantel over time in the WxM experiment, four analyses were conducted.

The ANOVA method was used to associate SNP with selection using the combination treatment of abamectin, oxfendazole and closantel. The comparison of the "drug", "unselected" and "mix" samples identified 27 SNPs with significant difference between the unselected and drug-selected samples (Fig 28). In 15 out of the 27 SNPs, the drug selected samples show a significantly higher allele frequency of the Alt allele. In the remaining 12 SNPs, the drug selected samples show a significantly lower Alt allele frequency. This almost 50:50 ratio between lower and higher allele frequency in the drug selected samples illustrates that the frequency changes are not skewed toward fixation at 0 more than at 1 or vice versa. However, the loss of heterozygosity in the drug

selected samples over all loci is evident from the skewed allele frequency plot (Fig 22). A further analyses of treatment by generation change in the allele frequency can be observed for several SNPs (Fig 29). Most notably the two BTUB1 mutations are both indicative of the generation number and treatment. These SNPs are therefore very important in the SNP-chip to detect selection due to the combination of drugs used.

The Fst by locus analysis for the WxM experiment was conducted but no SNPs were found to be associated with the combination drench selection using this method.

Fig 28. The 27 SNPs with significant differences between the SEL and UNS lines in the WxM dataset. **Figure unavailable in public access of this report due to IP.**

Fig 29. Minor allele frequency distribution by generation and treatment for most informative SNPs following the DAPC and PCA. **Figure unavailable in public access of this report due to IP.**

The principal component analysis (Fig 30) results explain 70% of the variation between samples from the WxM experiment. The dendrogram (Fig 30) shows the drench selected samples from passages 3, 4 and 5 clustered together, but the passage 6 sample is separated into a different cluster. A total of 5 SNPs are the most informative for identifying these 2 clusters (a second BTUB1 SNP is also implicated). Three SNP are also highly informative for defining the cluster containing the G1 and G2 samples with the UNS G3 and G4 samples. All these samples have an intermediate level of drench resistance compared to higher levels of resistance for the neighbouring clusters. The relationship between these SNP and response to selection with drenches is further illustrated in Fig 29. The response to selection can be seen where there are increasing or decreasing minor allele frequency values from G3_Drug through to G6_Drug, and the G3_Uns through to G6_Uns either have no pattern, or the opposite pattern to the Drug selected samples.

Fig 30. Dendrogram from the PCA clustering analysis along with most informative SNPs associated with specific generation clusters using the WxM data. **Figure unavailable in public access of this report due to IP.**

The RF method was used to interrogate the WxM SNP data associated with the treatment groups and the LDA results. These analyses used 70 SNPs, and could predict the UNS vs SEL categories with 72% accuracy. One locus, BTUB1, was significant for these categorical predictions (Fig 31). BTUB1 is a gene for which allelic variation is associated with benzimidazole resistance in many species of nematodes, and two of the three SNPs implicated from other research were included in our analysis (Fig 31A). Using EC50 values from the LDA, associations between abamectin resistance and thiabendazole resistance with SNPs were also sought using the RF method.

Fig. 31. (A) Categorical prediction of drench selected. **Figure unavailable in public access of this report due to IP.**

In summary from the WxM experiment, 31 SNPs (Table 13) were associated with oxfendazole, abamectin and/or closantel selection, four of these were significant using two separate statistical methodologies, and the two BTUB1 SNPs were significant using three methods. Although some of these SNP could have a causative relationship with drench resistance, it is better to consider these simply as associated with resistance in the Wallangra2003 population, as some likely are associated with resistance for reasons other than causation, such as genomic linkage with causative SNP or with genetic modifiers that enhance survival or resistant genotypes. The WxM experiment is also complicated in comparison to BxM due to the multi-drug selection, and the use of closantel in that combination of drenches. The categorical (SEL vs UNS) analyses are the only ones where closantel resistance is considered, whereas resistance to abamectin and benzimidazole compounds could be checked against methods which included EC50 values from the LDAs.

Table 13. SNP associated with selection with drench selection from analysis of the WxM experiment. **Table unavailable in public access of this report due to IP.**

Because the RF methodology is able to incorporate the EC50 data from the LDAs, we were able to use the data from the WxM experiment to look for associations between SNP and the other anthelmintics not used for selection. This was possible for levamisole as there was variation observed between samples in their survival response, but not for monepantel as there was little phenotypic variation in response to monepantel in the LDAs conducted using the WxM samples (Fig 20).

Fig. 32. Prediction of EC50 values for levamisole from LDA data. **Figure unavailable in public access of this report due to IP.**

4.3.2.4 SNPs identified in BxM, WxM and spatial differentiation experiments

The RF method identified six SNPs from the BxM experiment that predicted monepantel and levamisole, and seven SNPs from the WxM experiment that predicted benzimidazole, levamisole, abamectin and overall multiple drug selection (Table 14). Only two of the SNPs could be used individually to predict drug selection, the remaining SNPs need to be used in specific combinations.

Table 14. SNPs identity with ALT allele frequency predictive value range for a specific drug EC50 mode of action from the RandomForest method. SNPs with the same letter need to be taken as a combination to accurately predict corresponding drug resistance. **Figure unavailable in public access of this report due to IP.**

The WxM experiment allowed identification of 31 SNPs (Table 13), of which 19% (6 SNPs) were identified from two or more methods and 16% (6 SNPs) were also identified using the BxM experiment. The BxM experiment identified 14 SNPs (Table 12) associated with monepantel selection with an additional three SNPs (Table 14) associated with levamisole susceptibility. Five of those SNPs (30%) were identified with two or more methods. The 34 spatial population dataset identified 10 SNPs (Fig 13) that identified specific groups of populations.

5 Discussion

5.1 Genetic Drive Chip usefulness

5.1.1 SNP discovery process

This study has identified just over 13 million high confidence SNPs in the *H. contortus* or BPW genome. We used 0.0025% of these SNPs focusing on bi-allelic ones, in the research conducted. Future research which includes genetic characterization of *H. contortus* could use more of these identified SNP.

The SNP discovery pipeline used in this study (Fig 3) has proven successful; it is suitable to low quality and/or multiple reference genome assemblies, only a low percentage (2%) of SNPs did not amplify and the estimated allele frequency divergence between that predicted from sequence alignment and that predicted from Sequenom allelotyping was negligible (mean difference of 0.03). Primer design for the allelotyping is sensitive to SNPs close to the SNP to be assayed. Because we had catalogued so many SNP, we were able to concentrate upon SNP which had fewer variants nearby in the genome. This strategy has proven successful with the low drop-out rate we observed.

5.1.2 Quality-control steps

The quality control steps used in this study were necessary to eliminate SNPs that were not reliable for estimating allele frequencies from pooled samples, not reliable for identifying mixed population samples and/or were not *H. contortus* specific. In addition, SNPs were required to perform within multiplex reactions to minimise costs both for the research and subsequent commercial use on farms. The quality control steps which removed most candidate SNP were the multiplexing and the species-specific criteria. In field collected samples, each SNP will be genotyped in duplicate which will allow quantification of the fail rate of genotyping and suggest any issues with co-infecting species; although we will utilise species estimation by microscopy at the collecting laboratory to guard against this possibility in the first instance.

Given the genetic profile of two or more BPW populations, the genetic drive chip can identify the levels of admixture (0%, 25%, 50% or 75%) between populations. This capacity enabled us to consider detection of quarantine failure or other sources of between-property transfer of parasites and our ability to use the Genetic Drive Chip to distinguish these phenomena from selection for drench resistance within the property.

We attempted to identify factors which might predict failure of quality control for the SNP which we analysed in the project. None of the factors evaluated, such as sequencing coverage, loci duplication in either reference genome, frequency uncertainty values, low difference between ALT frequency of WAL and MCM, scaffold bias, homozygote to heterozygote ratio, or transversion vs transition (see B.AHE.0315 report 1 submitted on Dec 15th 2017) could successfully predict which SNPs failed or passed the quality control regime. This demonstrates the requirement for laboratory-based quality control steps rather than relying solely upon bioinformatics.

5.1.3 Population differentiation and migration detection using the genetic drive chip

The genetic drive chip was able to distinguish the 34 BPW populations analysed. Differentiation between the field and lab-derived populations, and between the various field populations was possible using four methodologies: PCA, hierarchical clustering, Shannon H index and multi-locus genotype network construction.

The first two dimensions of the PCA explained only 24% of the total variation in this dataset which indicates the complexity of this dataset's genetic profile. The collection was made across Australia (and South Africa) in 9 different regions spanning multiple climate zones, elevation and most importantly from varying sheep enterprises with different field history and management practices. It is most likely that genetic characteristics of these populations are the result of multiple factors beyond drench resistance and geographical location. However, drench resistance profile explained the genetic variation better than geographic location, indicating the potential of the SNP-chip to classify populations based on resistance status. Ten specific SNPs were identified (Fig 13) that can differentiate resistant and susceptible populations, without specifically focussing on any particular drench groups. Interestingly, one of the SNP which is located in the monepantel-resistant mutant gene (KY983243.1 - SNP s39045p6330) identifies a unique cluster of Invetus populations with unknown field history. Additionally, the non-transcribed (intergenic or intron) SNP ISEsc159p30845 identifies a cluster of five populations of which four (Goulburn, Wongarbon, Cooma and Bombala) are potentially resistant to macrocyclic lactones, closantel and levamisole. This illustrates the potential for the PCA method to use SNPs to identify drug resistance phenotypes of unknown samples based on cluster categorization.

The significant differences in allelic diversity between populations using the Shannon H index method indicate the potential of this statistic for use in population genetic differentiation. The method was especially useful for distinguishing highly inbred populations (lab-derived) and genetically isolated populations (e.g. South African). This means that the Shannon H index method would be best suited for identifying the incursion of novel, and potentially pathogenic, BPW strain from overseas.

The multi-locus heterozygosity value comparison did not show any differences between 23 of the populations. This was explainable as the allele frequency of the SNPs chosen varied between 0-1 and

the summation of these overall SNPs can hide differences in individual SNPs. The value of this heterozygosity metric is as a measure of background genetic diversity against which the allelic frequency of specific SNPs can be compared to identify signatures of selection.

In summary, the PCA method has proven to be the most likely candidate for future development of the Genetic Drive Chip for use in commercial sheep enterprises, but the other methods have niche roles which could also be important in future research

5.1.4 Monitoring drug resistance using the genetic drive chip

We conducted multiple drench (WxM trial – abamectin, closantel and oxfendazole) and single drench (BxM trial – monepantel) selection trial experiments in triplicate. These experiments involved the mixing of resistant and susceptible isolates of BPW, subsequent random mating between and within the isolates in sheep over two passages (generations 1 and 2), and four more passages of the parasites with (SEL) or without (UNS) selection with drenches. The UNS populations were used as controls against which the SEL populations were evaluated. We used both *in vivo* (FECRT) and *in vitro* (LDA) methods to quantify resistance levels for each generation of the trials. DNA was extracted from individual adults and pools of larvae from each of the generations in the WxM and BxM trials, and used for allelotyping with the Genetic Drive Chip.

5.1.4.1 Detecting overall changes in BPW strain genetics

The comparison of WxM and BxM generations using SNPs revealed clear differentiation patterns based on the PCA and the subsequent PCA clustering analyses grouped the UNS and SEL samples in different clusters for WxM. The allele frequency distribution of the WxM samples also showed marked changes in allele frequencies where in SEL samples, the number of SNPs with intermediate allele frequency values declined over repeated drench selection passages indicating an increased level of allele fixation. Allele fixation is often associated with increased selection pressure and this is the most likely explanation in our study. Allele fixation can also happen by random genic drift, however, this is less likely over a low number of generations, especially when the sampling between passages was thousands of individuals as in our experiments.

The pairwise Fst method identified significant differences between the SEL samples from the BxM and WxM lines and their parental isolates. The Fst differences between the UNS and parental samples showed a lesser differentiation. This indicates the potential to detect the absence or presence of drug selection pressure. The detection of a strong drug selection pressure (e.g multiple drugs as in the WxM) can be made much earlier and with more statistical significance than selection using a single drug.

Comparison of populations to each other using a large set of SNPs is useful for quantifying the degree of genetic differentiation between populations. This genetic differentiation, along with field history data, can be associated with a change in drug efficacy or associated with other factors such as migration leading to admixed populations or other parasite management strategies.

For future use of the BPW genetic drive chip to assess overall genetic differences in strains or by years, the use of PCA and PCA clustering method, allele frequency distribution spectrum method and

the pairwise Fst method are recommended. These three methods showed observable and significant population genetic profile differences.

5.1.4.2 Detecting specific SNPs with predictive value for drench resistance

Comparing results between the two selection trials and the population differentiation work, five SNP appeared to have associations in multiple investigations (Table 12, 13, 14). One of these SNP s39045p6330, was implicated in all three of the investigations, in the BxM and WxM trials distinguishing SEL populations from UNS populations, and in the population differentiation study distinguishing one cluster of anonymous Invetus samples from another. The SNP is located within an exon of a gene described as "Neurotransmitter-gated ion-channel transmembrane domain", but further work would be necessary to establish whether this is a causative relationship. For the intended purpose of monitoring BPW within sheep enterprises over time, this SNP will be useful.

Four other SNP have been implicated by more than one of our investigations (see table 14). These SNPs are associated with selection for drench resistance, but not specifically for particular drenches or compounds. These four SNP will also be useful for monitoring BPW within sheep enterprises over time.

The analysis combining the results of LDA tests with SNP and utilising the RF statistical method, allowed us to search for SNP associated with specific drench resistance for benzimidazoles, macrocyclic lactones, monepantel and levamisole. Eleven SNPs (Table 14) were identified with these associations across the two selection experiments. Two SNP in the BTUB1 gene were associated with benzimidazole resistance and these had both been implicated in resistance to these drenches prior to this work. Three SNPs were identified which are associated with resistance to abamectin, one with monepantel and one with resistance to both abamectin and monepantel. Four other SNPs were found to have an association with levamisole resistance. These results confirm the usefulness of these eleven SNPs, but care should be taken with the interpretation of these results. Further work would be required to establish causative relationships between these SNP and drench resistance, however the associations will be useful for monitoring BPW populations with a known history in the near future.

Our investigations have discovered 51 SNPs associated with drench resistance at some level and these SNPs should be included within a future Genetic Drive Chip tool for sheep producers to use. However, because of the likelihood of differing mechanisms of resistance in different populations and also the "background" genetic diversity within and between BPW populations, the inclusion of more SNP which have no known association with drench resistance is advisable to make the Genetic Drive Chip tool useful.

5.2 Commercialization plan

The current diagnostic test used to detect drench resistance in the Australian sheep industry is the faecal egg count reduction test (FECRT). This test takes weeks to perform, is expensive ($587–$875) and its adoption rate is very low (<2%). This low adoption rate is probably due to a lack of producer understanding of the test's value as well as the undesirable cost and labour input. We aim to increase the adoption of testing for drench resistance in the sheep industry. Although very few

producers (<2%) currently undertake drench resistance testing, the adoption of worm monitoring of flocks (WEC) has had a higher adoption rate of between 10 and 20%. The BPW genetic drive chip proposed market price of $250 is much less than FECRT, and the labour input required is the same as for a standard flock monitor WEC test. A flock monitor WEC and BPW genetic drive chip can be can be undertaken using the same sample, meaning that no additional labour would be required for producers who have already adopted flock monitor WEC. The proposed price of the BPW genetic drive chip allows for a modest mark-up for the parties involved in producing the test result, it exceeds the base cost of the labour, consumables and overheads required to produce the test result.

A medium estimate of the cost of parasitism in Australian sheep is $12/hd/yr (MLA report B.AHE.0010), but the cost of BPW infection is likely higher. Ineffective drenching will cost the producer both the price per animal of the ineffective drench and the loss of production due to incomplete parasite removal. Where BPW causes a death, the value of the animal is lost in addition to the other losses. Conservatively, if four drenches per year were required to control parasitism, this $12 cost/hd/yr would consist of $1.20 to $3.20 for purchasing drench and approximately $9.00 in lost production without considering farm labour. An ineffective drench with efficacy of 70% might necessitate increasing the frequency of drenching by 25% to five times per year, and would increase the production losses by at least a similar amount, so that the cost of parasitism would rise from $9.00 to $11.25. In total, an additional cost per head of $2.55–$3.05. The best way to prevent this outcome is to use regular drench resistance testing and select the drenches proven to be effective by the test. Given these figures, an estimate can be made of the flock size necessary so that the cost of drench resistance testing will be less than the additional savings created by using effective drenches. The figure below provides a comparison of this for FECRT across the range of current prices and the potential DNA-based test (Fig 33). Using FECRT a flock size of greater than 200 would be needed to break even at the lowest commercial price, whereas exceeding 100 animals would give a break even outcome with the DNA-based test.

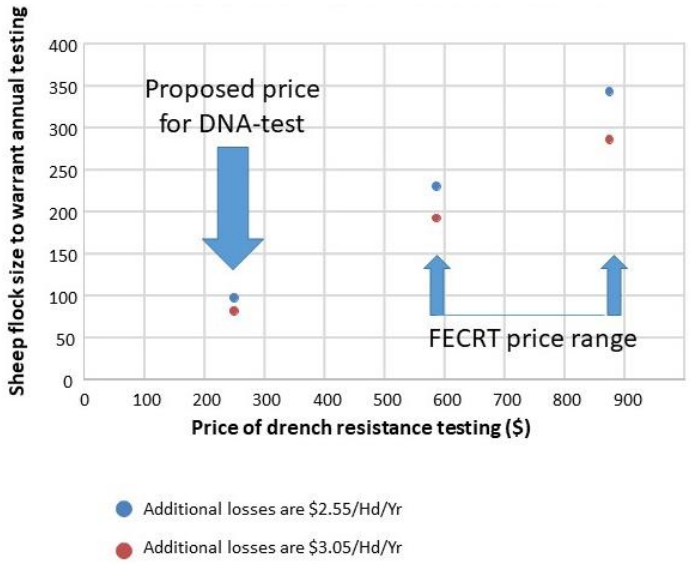

Fig 33. Marginal flock sizes necessary to break even given the two values of increased losses due to resistance and three prices for drench resistance testing. DNA-test in this figure refers to the predicted retail price for the BPW genetic drive chip.

The BPW SNP-chip requires cultured parasite larvae from a typical flock monitoring faecal collection on farm. The sample is submitted to a parasite diagnostic lab in the same way in which samples are

submitted for WEC. Producers will receive a WEC estimate from the lab in the same way they do now, and also an estimate of the proportion of larvae from the different parasite species after culturing the faeces and visual inspection as they do now. The new tests are conducted using some of these cultured larvae which are sent to a genotyping lab for analysis. At the genotyping lab, DNA is extracted from the worms and analysed; around 180 different DNA tests are conducted simultaneously in the test. The DNA results are compared to results from a database of other samples and importantly, also compared to samples from the same sheep operation which had been submitted in previous years. Results will indicate the likelihood of resistance against the various classes of drenches and also a general report showing how successful parasite control has been since the previous test.

To achieve this commercialization plan, a field trial of the Barber's pole worm genetic drive chip is necessary (see section 5.4.1).

5.3 Delivery on initial project objectives

The initial project objectives were;

1) Select Single Nucleotide Polymorphism (SNP) and insertion/deletion (InDel) markers from appropriate genome regions for the chip.

   This objective was met successfully with the selection of 334 SNPs across Barber's pole worm genome of which 197 successfully passed the quality control steps. These SNPs as a set can be used to assess genetic characteristics in this nematode species such as drench resistance.

2) Use the chip to analyse samples from five geographically separated BPW isolates with different chemical resistance profiles, to distinguish populations from each other.

   This objective was met using 34 field populations collected from 9 regions across Australia. Populations were differentiated using the genetic drive chip SNPs and multiple statistical methods. Of the populations sampled, 15 had a known drug resistance phenotype. Drench resistance rather than geographical origin was shown to be the best explanatory variable for the differing genetic profiles of the samples analysed.

3) Use the chip to track genetic change following chemical selection pressure

   We generated drug selected and unselected lines for two different drench combinations and tracked genetic change over generations (passages) using the genetic drive chip SNPs. The changes in drug resistance status were accompanied by significant changes in allele frequency and genetic profile. Specific SNPs were associated with significant differences between selected and unselected lines and the prediction of drug resistance level. Four different statistical methods were used to associate SNPs with drench selection and with resistance against specific drench compounds.

5.4 Future research recommendations

Four recommendations are made to MLA. These are ordered by priority and importance.

5.4.1 Pre-commercialization field trial of BPW genetic drive chip

To be ready for commercial release we advocate a field trial of the BPW genetic drive chip. The field trial should test the supply chain from the producer obtaining the sample, through the two laboratory stages and returning recommendations back to the producer. The trial should involve 100 to 160 sheep production enterprises, sampled from across the BPW endemic zone, and including some representative participants from regions where sporadic BPW outbreaks occur. Samples will be submitted for allelotyping analysis from each participating enterprise each year for three consecutive years, this will test the ability of the supply chain to meet demand, and will expand the statistical repertoire of the tests for detecting drench resistance across multiple compounds and multiple geographic regions. The necessary association of the work with careful within-enterprise record keeping will have additional benefits beyond the goal of developing the BPW genetic drive chip. To prove the usefulness of the test, FECRT will need to be conducted on the same properties to allow a direct comparison. Producers who are involved with the work will have input into the outcome of the trial, especially regarding the way in which results are interpreted and presented for practical use within the sheep enterprise. Consultation with producers will occur to help create a user friendly software interface to deliver test outcomes. The project will involve extensive communication with industry stakeholders and collaboration with commercial partners to make the tool applicable to the end users and ready to commercialise.

This project has been submitted to AWI for funding (complete application in Appendix 9.9) as the 2019 MLA funding priorities did not include research aiming to improve livestock health.

5.4.2 Addition of other parasite species to the genetic drive chip

In this project we have demonstrated the usefulness of the genetic drive chip methodology for monitoring and predicting drench resistance and other genetics characteristics of BPW. This has been a useful demonstration of the principle, and consideration should now be given toward expanding the tool to deal with other organisms of concern for sheep and cattle producers. There are two main reasons we believe this is both important and feasible.

Firstly, the suite of GIN, and other GI tract parasites of importance to the sheep and cattle industries comprises at least *Teladorsagia circumcincta*, *Trichostrongylus colubriformis*, *Tr. vitrinus*, *Ostertagia ostertagi*, *H. placei* and *Fasciola hepatica* in addition to *H. contortus*. All these parasites can cause significant economic loss under certain circumstances, and there are both regions of Australia, and specific production systems, where one or more of these is more important than the others.

Secondly, there are significant economies of scale opportunities for SNP-based allelotyping. Although the Sequenom methodology we used in this project can deliver up to a few hundred SNP genotypes with reasonable cost, the hybridisation-based alternative methods can deliver thousands of tests for very little additional cost once developed. SNP assays for very many species could be assembled in the same set and allelotyped for all samples, achieving a significantly improved test range with a much wider market reach. With some forethought, there may even be wider market

opportunities if other key organisms of importance were included beyond GI tract parasites, to bacterial organisms, ectoparasites, or even beyond sheep and cattle health to pasture health or other agricultural applications.

The research pipeline needed to deliver a tool similar to the BPW genetic drive chip would be similar to the investment made in the current project if the species had a genome sequence available, even in a rudimentary form, and an available collection of isolates from different populations. In this regard the most near to market parasites would be the GI tract species *Te. circumcincta*, *F. hepatica* and *H. placei*, and the ectoparasites *Lucilia cuprina* and *Rhipicephalus microplus*. Research to include other species with less developed genomic information would take longer, and need to include some genome sequencing or re-sequencing work

5.4.3 Drug hypersensitivity and other interactions between anthelmintics

The *in vivo* drug selection experiment and subsequent target and non-target drug resistance characterization allowed to identify patterns of cross-resistance associated with resistance to specific drugs. This was not the focus of this study, however, the findings have implication for the drug resistance management of BPW and its diagnostics.

The observation that an increase in resistance to one drug was associated with a decrease in resistance to another drug mode of action is the core of the pesticide hypersensitivity hypothesis (to be distinguished from the drug hypersensitivity reaction in humans). This phenomenon is explained by the genetic trade-off, essentially a fitness cost, of resistance to one drug/pesticide. This phenomenon has been observed for insecticide resistance in moths between imidacloprid and methomyl (Abbas *et al* 2012). Pesticide resistance management depends closely on the presence of fitness cost and sensitivity to drugs (Hall et al 2004), therefore an in depth investigation of cross-resistance patterns between drugs, associated with genetic resistance mechanism, could yield significant new information which could be used in drench resistance management plans in BPW.

During B.AHE.0315 we observed a significantly slower selection for drug resistance to the combination of closantel, oxfendazole and abamectin, compared to the single active monepantel. We observed longer times (+14 days) post-drench for eggs to be produced and observed fewer and smaller adult individuals in the host abomasum for the multiple drug treatments compared to the single drug treatments. This life cycle trade-off between multiple drug resistance and adult maturity and establishment should be investigated further. Changes in life cycle traits could interact with non-chemical control strategies to minimize multiple drug resistant populations. These observations confirm the work of others suggesting that combination products are less prone to selection for resistance compared to single actives.

We also observed the association of a single SNP with resistance to both abamectin and monepantel which is suggestive of a non-specific resistance mode of action. Additionally, some of the SNP with drench resistance associations identified were within genes which have been implicated in non-specific modes of action, such as cytochrome P450 genes (2 SNP) and ABC transporter genes (1 SNP). Further research on the genetic interactions between resistance against currently available anthelmintic compounds could be crucial for predicting the outcomes for new combination drugs, and producing more sophisticated guidelines for drench rotation for sheep producers.

The implications of these cross-resistance patterns could inform the management solutions resulting from the genetic drive chip genetic profiling. This highlights the importance of obtaining the genetic profile using the genetic chip drive from multiple BPW strains showing various pattern of drug cross-resistance in order to assign a genetic profile with as many drug resistance pattern as possible. This association map would allow the genetic drive chip to quickly and reliably classify unknown samples into drug resistance categories.

5.4.4 The *in vivo* selection trial

The in-vivo drug selection experiments generated in this study represent a unique opportunity to study the microevolution of drug resistance selection to different drugs. Not many *in vivo* drug selection experiments have been performed worldwide on BPW, making the samples, preserved live larvae and the dataset unique and valuable. An in-depth analysis of selected and unselected replicates at each generation in the two lines using next generation sequencing (genomic and/or transcriptomic) could identify non-SNP signatures of drug selection which could become a new diagnostic. Genomic and transcriptomic analysis could also identify loci strongly associated with specific drug resistance indicating the mechanisms of resistance and therefore informing drench resistance management. Genomic and transcriptomic analysis could also identify new drug targets that would be antagonistic to drug resistance in BPW. Two of the drug resistances used in these experiments have been substantially less studied than others; monepantel because it has not been available as long as the others, and closantel because it is not used in Europe and North America. Insights into drench resistance against these two compounds would be useful both scientifically and for practical applications. Such work may lend itself to graduate student project(s).

6 Conclusions/recommendations

The BPW Genetic drive chip has been developed and tested against multiple and diverse samples of BPW. It can detect the progress of selection towards drench resistance and differences between isolates of BPW. This new diagnostic can be delivered for sheep producers to use after a field evaluation and supply chain testing project which we recommend should be undertaken over three years, and can begin from July, 2019.

The project has provided additional insights into other aspects of drench resistance which could also be the focus of future research.

6.1 Future R&D

The BPW genetic drive chip can be delivered for sheep producers to use after a field evaluation and supply chain testing project which we recommend should be undertaken over three years, and can begin from July, 2019.

Other research should be considered as a consequence of undertaking the work described here. See section 5.4 for details.

6.2 Extension activities

In order to begin the commercialization of the BPW genetic drive chip, we recommend a field evaluation and supply chain testing project. See section 5.4 and appendix 9.9 for details.

7 Key messages

Six key messages emerge from this work:

7- The BPW genetic drive chip is a new drug resistance diagnostic which utilises a panel of over 100 SNP assays.
8- The project has reinforced the view that the complexity of drench resistance genetics does not allow simple, single gene test for resistance diagnostics.
9- The project demonstrated that the BPW genetic drive chip can be used with field-derived populations to follow drench resistance.
10- The BPW genetic drive chip can be delivered to the industry at a retail price of between one quarter and one third of FECRT, and a substantial decrease in labour needed on farm compared to FECRT.
11- A three year field evaluation and supply chain testing project, conducted in close collaboration with more than 100 producers is the recommended next step towards commercialising the BPW genetic drive chip.
12- The genetic drive chip can be made to incorporate many additional species of importance for sheep and cattle producers with very little increase in the retail price of the diagnostic test. Further research will be required to incorporate more species.

8 Bibliography

Abbas N, Shad SA, Razaq M (2012) Fitness cost, cross resistance and realized heritability of resistance to imidacloprid in Spodoptera litura (Lepidoptera: Noctuidae). *Pesticide Chem and Physiology* **103**, 181-188.

Boitard S, Schlotterer C, Nolte V, Pandey RV, Futschik A (2012) Detecting Selective Sweeps from Pooled Next-Generation Sequencing Samples. *Molecular Biology and Evolution* **29**, 2177–2186. doi:10.1093/molbev/mss090.

Breiman L (2001) Random forests. *Machine Learning* **45**, 5-32.

Bryant D, Moulton V (2003) Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular Biology and Evolution* **21**, 255–265. doi:10.1093/molbev/msh018.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491–498. doi:10.1038/ng.806.

Emery DL, Hunt PW, and Le Jambre LF (2016) Haemonchus Contortus: The Then and Now, and Where to from Here?. *International Journal for Parasitology* **46**, 755–69. https://doi.org/10.1016/j.ijpara.2016.07.001.

Futschik A, Schlötterer C (2010) The Next Generation of Molecular Markers From Massively Parallel Sequencing of Pooled DNA Samples. *Genetics* **186**, 207–218. doi:10.1534/genetics.110.114397.

Gill JH, Redwin JM, van Wyk, JA, Lacey E (1995) Avermectin inhibition of larval development in Haemonchus contortus - effects of ivermectin resistance. *Int. J. Parasitol*. **25**, 463-470.

Hall R, Gubbins S, Gilligan CA (2004) Invasion of drug and pesticide resistance is determined by trade off between treatment efficacy and relative fitness. *Bulletin of Mathematical Biology* **66**, 825-840.

Hunt PW, Knox MR, Le Jambre LF, McNally J, Anderson LJ (2008) Genetic and phenotypic differences between isolates of Haemonchus contortus in Australia. *International Journal for Parasitology* **38**, 885–900. doi:10.1016/j.ijpara.2007.11.001.

Kassambara A (2017) Practical guide to Principal component method in R. Edition 1. Multivariate Analysis II. STHPA Publishing.

Kotze AC, O'Grady J, Emms J, Toovey AF, Hughes S, Jessop P, Bennell M, Vercoe PE, Revell DK (2009) Exploring the anthelmintic properties of Australian native shrubs with respect to their potential role in livestock grazing systems. *Parasitology* **136**, 1065. doi:10.1017/S0031182009006386.

Kotze AC, and Prichard RK (2016). Anthelmintic Resistance in Haemonchus Contortus. In *Advances in Parasitology* **93**, 397–428. https://doi.org/10.1016/bs.apar.2016.02.012.

Kotze AC, Ruffell A, Lamb J, Elliott TP (2018) Response of drug-susceptible and -resistant Haemonchus contortus larvae to monepantel and abamectin alone or in combination in vitro. *Veterinary Parasitology* **249**, 57–62. doi:10.1016/j.vetpar.2017.11.007.

Lacey E, Redwin, JM, Gill JH, Demargheritit VM, and Waller PJ (1990) A larval development assay for the simultaneous detection of broad spectrum anthelmintic resistance. In: Boray, J.C., Martin, P.J. Roush, R.T. (Eds.), Resistance of Parasites to Antiparasitic Drugs. MSDAGVET, Rahway, NJ, pp 177-184.

Laing R, Kikuchi T, Martinelli A, Tsai IJ, Beech RN, Redman E, Holroyd N, Bartley DJ, Beasley H, Britton C, others (2013) The genome and transcriptome of Haemonchus contortus, a key model parasite for drug and vaccine discovery. *Genome biology* **14**, R88.

Lane J, Jubb T, Shepherd R, Webb-Ware J, Fordyce G (2015) Priority list of endemic diseases for the red meat industries. Meat and Livestock Australia, B.AHE.0010. (Sydney)

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079. doi:10.1093/bioinformatics/btp352.

Love S, Neilson F, Biddle A, McKinnon R (2003) Moxidectin-resistant Haemonchus contortus in sheep in northern NSW. *Australian Veterinary Journal* **81**, 359–360.

Love S (2011) Drench resistance and sheep worm control. NSW Department of Primary Industries, Primefact **478**, second edition.

Lynch M, Bost D, Wilson S, Maruki T, Harrison S (2014) Population-Genetic Inference from Pooled-Sequencing Data. *Genome Biology and Evolution* **6**, 1210–1218. doi:10.1093/gbe/evu085.

Lyndal-Murphy M, Ehrlich WK, and Mayer DG (2014). Anthelmintic Resistance in Ovine Gastrointestinal Nematodes in Inland Southern Queensland. *Australian Veterinary Journal* **92**, 415–20. https://doi.org/10.1111/avj.12250.

Meat and Livestock Australia (2016) Australia's sheep industry. Meat and Livestock Australia.

Raza A, Lamb J, Chambers M, Hunt PW, Kotze AC (2016) Larval development assays reveal the presence of sub-populations showing high- and low-level resistance in a monepantel (Zolvix®)-resistant isolate of Haemonchus contortus. *Vet Parasitol*. **220**, 77-82

R Development Core team (2017) R: A language and environment for statistical computing. (R Foundation for Statistical Computing: Vienna, Austria) http://www.R-project.org.

Rellstab C, Zoller S, Tedder A, Gugerli F, Fischer MC (2013) Validation of SNP Allele Frequencies Determined by Pooled Next-Generation Sequencing in Natural Populations of a Non-Model Plant Species (GA Marais, Ed.). *PLoS ONE* **8**, e80422. doi:10.1371/journal.pone.0080422.

Ruffell A, Raza A, Elliott TP, Kotze AC (2018) The use of the larval development assay for predicting the in vivo efficacy of levamisole against Haemonchus contortus and Trichostrongylus colubriformis. *Vet Parasitol*. **260**, 6-11

Schwarz EM, Korhonen PK, Campbell BE, Young ND, Jex AR, Jabbar A, Hall RS, Mondal A, Howe AC, Pell J, others (2013) The genome and developmental transcriptome of the strongylid nematode Haemonchus contortus. *Genome biology* **14**, R89.

Shannon CE, Weaver W (1949) The mathematical theory of communication. Illinois: University of Illinois press.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA (2013) From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline: The Genome Analysis Toolkit Best Practices Pipeline. 'Curr. Protoc. Bioinforma.' (Eds A Bateman, WR Pearson, LD Stein, GD Stormo, JR Yates) p. 11.10.1-11.10.33. (John Wiley & Sons, Inc.: Hoboken, NJ, USA) doi:10.1002/0471250953.bi1110s43.

9 Appendix

9.1 SNP discovery pipeline UNIX script

Appendix 9.1 unavailable in public access of this report due to IP.

9.2 Summary table of 334 SNPs of the genetic drive chip

This appendix is provided as a separate excel file (BAHE0315_Appendix9.2_SNPsummaryTable.xls) and not included in this document due to its size.

The sequences of primers used for amplification and primer extension and the sets which go together to form each "plex" is sensitive intellectual property which cannot be protected except as a "trade secret".

Appendix 9.2 unavailable in public access of this report due to IP.

9.3 SNP quality control R script

Appendix 9.3 unavailable in public access of this report due to IP.

9.4 SNP population genetics R script

Figure unavailable in public access of this report due to IP.

9.5 Table of individual NGS library processing statistics

	Reads from Illumina	Trimming reduction %	% aligned to ref genome	filtering reduction%	Read duplication %	mean coverage	> min cov 10
Aligned to the ISE genome							
mcm8264	44634640	7.55	86	58	7.8	4.73	14
mcm8752	66334112	5.27	88	56	8.6	7.53	33
mcm8756	52579388	4.28	89	57	2.3	6.35	26
mcm8801	54530914	7.04	85	59	8.4	5.57	20
mcm8850	47022620	5.79	87	58	5.5	5.17	17

mcm9488	55950044	4.28	89	70	4.5	6.68	29
ALL.MCM	321051718	5.7	87	60	6	36.02	80
wal8229	56142084	5.44	88	55	6.40	6.41	29
wal8799	50889590	3.45	90	55	2.90	6.77	27
wal8837	48142322	4.31	89	57	3.50	5.82	22
wal8840	68697396	5.61	87	56	6.30	7.94	36
wal9225	62892848	5.79	88	56	8.10	7.26	32
ALL.WAL	286764240	4.92	88	56	5.40	34.2	82

Aligned to the McM genome						
mcm8264		86	72	7.9	3	6
mcm8752		88	70	8.5	4.81	15
mcm8756		89	71	2.6	4.04	11
mcm8801		85	72	8.5	3.52	8
mcm8850		87	72	5.5	3.24	7
mcm9488		89	56	4.7	4.31	13
ALL.MCM		88	69	6.2	22.92	73
wal8229		89	70	6.6	4.24	13
wal8799		90	70	3.1	3.95	11
wal8837		90	71	3.6	3.6	9
wal8840		88	71	6.3	4.87	16
wal9225		88	70	8.1	4.57	14
ALL.WAL		89	70	5.5	21.23	72

9.6 Larval development assay raw data

This appendix is provided as a separate excel file (BAHE0315_Appendix9.6_LDARawData.xls) and not included in this document due to its size.

Unpublished data made unavailable to the public.

Appendix 9.6 unavailable in public access of this report due to IP.

9.7 Table of drug EC50 values and EC50 pairwise significant differences

Appendix 9.7. Effective dose inhibiting development of 50% of the individuals (EC50) for each drug tested with 95% confidence intervals from *H.contortus* parental isolates to generation four in the in vivo drug selectin trial. TBZ, thiabendazole, ABA; abamectine, MPTL; monepantel, LEV; levamisole.

Drug	Isolate /Line	G0	G1	G2	G3	G4	G5	G6
TBZ	McMaster1931	0.01218 (0.009-0.001)						
	Wallangra2003ABA	0.12 (0.10-0.15)						
	Wallangra2003OXF	0.77 (0.62-0.96)						
	Bundarra2014	7.67 (6.6-8.7)						
	BxM DrugSel				0.233 (0.186-0.293)	0.538 (0.414-0.698)	0.5046 (0.356-0.708)	0.7995 (0.6586-0.9687)
	BxM Uns		0.102 (0.62-0.142)	0.180 (0.11-0.25)				0.1751 (0.1547-0.1982)
	WxM DrugSel				0.024 (0.0213-0.0271)	0.2483 (0.211-0.291)	0.2393 (0.2096-0.2735)	0.3877 (0.3463-0.4337)
	WxM Uns		0.033 (0.030-0.036)	0.032 (0.029-0.034)				0.0211 (0.0195-0.0229)
ABA	McMaster1931	0.00011 (0.00010-0.00014)						

	Wallangra2003ABA	0.00079 (0.00070-0.00066)						
	Wallangra2003OXF	0.00050 (0.00038-0.00066)						
	Bundarra2014	0.00071 (0.00062-0.00082)						
	BxM DrugSel				0.00044 (0.00039-0.00051)	0.00032 (0.00028-0.00037)	0.00045 (0.00040-0.00049)	0.000414 (0.00038-0.00044)
	BxM Uns		0.00056 (0.00052-0.00060)	0.00032 (0.00028-0.00035)				0.000437 (0.000414-0.000460)
	WxM DrugSel				0.00030 (0.00027-0.00034)	0.00045 (0.00037-0.00053)	0.000527 (0.000472 - 0.000586)	0.000711 (0.000663-0.000761)
	WxM Uns		0.00059 (0.00056-0.00062)	0.00051 (0.00047-0.00054)				0.00042 (0.00039-0.00045)
MPTL	McMaster1931	0.0036 (0.0030-0.0043)						
	Wallangra2003ABA	0.0072 (0.006-0.008)						
	Wallangra2003OXF	0.0023 (0.0018-0.0030)						
	Bundarra2014	2.52 (2.27-2.76)						
	BxM DrugSel				0.0067 (0.0053-0.0086)/4.08 (2.7-7.0)	0.0058 (0.0028-0.0121)/4.157 (3.454-5.003)	4.168 (3.319-5.152)	5.619 (5.114-6.126)
	BxM Uns		0.006 (0.004-0.007)/5.13 (3.21-15.41)	0.005 (0.003-0.0075)/9.4 (large)				0.00558 (0.00485-0.00636)/4.7555 (large)

	WxM DrugSel				0.0034 (0.0027-0.0043)	0.0031 (0.0029-0.0033)	0.004234 (0.00380-0.00470)	0.00452 (0.00435-0.00469)
	WxM Uns		0.0043 (0.0041-0.0045)	0.0029 (0.0027-0.0031)				0.00437 (0.00415-0.00459)
LEV	McMaster1931	0.00399 (0.0000079-0.0289)						
	Wallangra2003ABA x Wallangra2003OXF	0.561 (0.394-0.816)						
	Bundarra2014	0.2317 (0.203-0.2655)						
	BxM DrugSel					0.188 (0.168-0.209)	0.326 (0.277-0.378)	0.191 (0.177-0.206)
	BxM Uns			0.2092 (0.187-0.233)				0.239 (0.225-0.254)
	WxM DrugSel					0.1976 (0.155-0.255)	0.503 (0.380-0.672)	0.352 (0.306-0.406)
	WxM Uns			0.300 (0.248-0.400)				0.259 (0.230-0.291)

Significantly different generation pairwise comparison of EC50.

Line	Drug	pairwise population comparison	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value
WxM	ABA	NA	NA	NA	NA	NA
WxM	LEV	McMaster1931 vs. Wallangra2003	-0.5578	-0.655 to -0.4606	****	<0.0001
WxM	LEV	McMaster1931 vs. G2	-0.2962	-0.3892 to -0.2032	****	<0.0001
WxM	LEV	McMaster1931 vs. G4	-0.1936	-0.3002 to -0.08698	****	<0.0001

WxM	LEV	McMaster1931 vs. G5	-0.4992	-0.6015 to -0.3969	****	<0.0001
WxM	LEV	McMaster1931 vs. G6DRUG	-0.3488	-0.433 to -0.2646	****	<0.0001
WxM	LEV	McMaster1931 vs. G6UNS	-0.2555	-0.3403 to -0.1707	****	<0.0001
WxM	LEV	Wallangra2003 vs. G2	0.2616	0.1743 to 0.3489	****	<0.0001
WxM	LEV	Wallangra2003 vs. G4	0.3642	0.2625 to 0.4659	****	<0.0001
WxM	LEV	Wallangra2003 vs. G6DRUG	0.209	0.1312 to 0.2868	****	<0.0001
WxM	LEV	Wallangra2003 vs. G6UNS	0.3023	0.2238 to 0.3808	****	<0.0001
WxM	LEV	G2 vs. G4	0.1026	0.004928 to 0.2003	*	0.0323
WxM	LEV	G2 vs. G5	-0.203	-0.296 to -0.11	****	<0.0001
WxM	LEV	G4 vs. G5	-0.3056	-0.4122 to -0.199	****	<0.0001
WxM	LEV	G4 vs. G6DRUG	-0.1552	-0.2445 to -0.06585	****	<0.0001
WxM	LEV	G5 vs. G6DRUG	0.1504	0.06622 to 0.2346	****	<0.0001
WxM	LEV	G5 vs. G6UNS	0.2437	0.1589 to 0.3285	****	<0.0001
WxM	LEV	G6DRUG vs. G6UNS	0.0933	0.03164 to 0.155	***	0.0002
WxM	TBZ	McMaster1931 vs. Wallangra2003	-0.3006	-0.3224 to -0.2788	****	<0.0001
WxM	TBZ	McMaster1931 vs. G1	-0.02626	-0.04676 to -0.005765	**	0.0024
WxM	TBZ	McMaster1931 vs. G4	-0.2361	-0.2596 to -0.2126	****	<0.0001
WxM	TBZ	McMaster1931 vs. G5	-0.2271	-0.2502 to -0.204	****	<0.0001
WxM	TBZ	McMaster1931 vs. G6DRUG	-0.3755	-0.3965 to -0.3545	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G1	0.2744	0.2618 to 0.287	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G2	0.2832	0.2709 to 0.2955	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G3	0.2887	0.2707 to 0.3068	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G4	0.0645	0.04742 to 0.08158	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G5	0.0735	0.05696 to 0.09004	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G6DRUG	-0.0749	-0.08826 to -0.06154	****	<0.0001
WxM	TBZ	Wallangra2003 vs. G6UNS	0.2916	0.2782 to 0.3051	****	<0.0001
WxM	TBZ	G1 vs. G4	-0.2099	-0.2252 to -0.1945	****	<0.0001
WxM	TBZ	G1 vs. G5	-0.2009	-0.2156 to -0.1861	****	<0.0001
WxM	TBZ	G1 vs. G6DRUG	-0.3493	-0.3603 to -0.3382	****	<0.0001

WxM	TBZ	G1 vs. G6UNS	0.01727	0.006113 to 0.02843	****	<0.0001
WxM	TBZ	G2 vs. G4	-0.2187	-0.2338 to -0.2036	****	<0.0001
WxM	TBZ	G2 vs. G5	-0.2097	-0.2242 to -0.1952	****	<0.0001
WxM	TBZ	G2 vs. G6DRUG	-0.3581	-0.3689 to -0.3473	****	<0.0001
WxM	TBZ	G3 vs. G4	-0.2242	-0.2443 to -0.2042	****	<0.0001
WxM	TBZ	G3 vs. G5	-0.2152	-0.2348 to -0.1957	****	<0.0001
WxM	TBZ	G3 vs. G6DRUG	-0.3636	-0.3806 to -0.3466	****	<0.0001
WxM	TBZ	G4 vs. G6DRUG	-0.1394	-0.1554 to -0.1234	****	<0.0001
WxM	TBZ	G4 vs. G6UNS	0.2271	0.2111 to 0.2432	****	<0.0001
WxM	TBZ	G5 vs. G6DRUG	-0.1484	-0.1638 to -0.133	****	<0.0001
WxM	TBZ	G5 vs. G6UNS	0.2181	0.2027 to 0.2336	****	<0.0001
WxM	TBZ	G6DRUG vs. G6UNS	0.3665	0.3545 to 0.3785	****	<0.0001
WxM	MPTL	NA	NA	NA	NA	NA
BxM	ABA	NA	NA	NA	NA	NA
BxM	LEV	McMaster1931 vs. Bundarra2014	-0.2277	-0.3222 to -0.1332	****	<0.0001
BxM	LEV	McMaster1931 vs. G2	-0.2052	-0.2874 to -0.123	****	<0.0001
BxM	LEV	McMaster1931 vs. G4	-0.184	-0.2898 to -0.07819	****	<0.0001
BxM	LEV	McMaster1931 vs. G5	-0.3227	-0.4197 to -0.2257	****	<0.0001
BxM	LEV	McMaster1931 vs. G6drug	-0.1871	-0.2692 to -0.105	****	<0.0001
BxM	LEV	McMaster1931 vs. G6uns	-0.2355	-0.3181 to -0.1529	****	<0.0001
BxM	LEV	Bundarra2014 vs. G5	-0.095	-0.1883 to -0.001653	*	0.043
BxM	LEV	G2 vs. G5	-0.1175	-0.1983 to -0.03667	***	0.0004
BxM	LEV	G4 vs. G5	-0.1387	-0.2434 to -0.03395	**	0.0019
BxM	LEV	G5 vs. G6drug	0.1356	0.05488 to 0.2163	****	<0.0001
BxM	LEV	G5 vs. G6uns	0.0872	0.006005 to 0.1684	*	0.0261
BxM	TBZ	McMaster1931 vs. Bundarra2014	-7.664	-7.712 to -7.615	****	<0.0001
BxM	TBZ	McMaster1931 vs. G1	-0.3261	-0.3639 to -0.2883	****	<0.0001
BxM	TBZ	McMaster1931 vs. G2	-0.3477	-0.3856 to -0.3099	****	<0.0001

BxM	TBZ	McMaster1931 vs. G3	-0.2215	-0.264 to -0.179	****	<0.0001
BxM	TBZ	McMaster1931 vs. G4	-0.5259	-0.5696 to -0.4822	****	<0.0001
BxM	TBZ	McMaster1931 vs. G5	-0.4924	-0.5345 to -0.4503	****	<0.0001
BxM	TBZ	McMaster1931 vs. G6Drug	-0.7873	-0.8257 to -0.7489	****	<0.0001
BxM	TBZ	McMaster1931 vs. G6Uns	-0.1629	-0.201 to -0.1248	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G1	7.338	7.302 to 7.373	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G2	7.316	7.281 to 7.351	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G3	7.442	7.402 to 7.483	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G4	7.138	7.096 to 7.179	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G5	7.171	7.132 to 7.211	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G6Drug	6.877	6.841 to 6.912	****	<0.0001
BxM	TBZ	Bundarra2014 vs. G6Uns	7.501	7.465 to 7.536	****	<0.0001
BxM	TBZ	G1 vs. G2	-0.0216	-0.03931 to -0.003891	**	0.005
BxM	TBZ	G1 vs. G3	0.1046	0.07835 to 0.1309	****	<0.0001
BxM	TBZ	G1 vs. G4	-0.1998	-0.228 to -0.1716	****	<0.0001
BxM	TBZ	G1 vs. G5	-0.1663	-0.1919 to -0.1407	****	<0.0001
BxM	TBZ	G1 vs. G6Drug	-0.4612	-0.4801 to -0.4423	****	<0.0001
BxM	TBZ	G1 vs. G6Uns	0.1632	0.1449 to 0.1815	****	<0.0001
BxM	TBZ	G2 vs. G3	0.1262	0.09992 to 0.1525	****	<0.0001
BxM	TBZ	G2 vs. G4	-0.1782	-0.2064 to -0.15	****	<0.0001
BxM	TBZ	G2 vs. G5	-0.1447	-0.1703 to -0.1191	****	<0.0001
BxM	TBZ	G2 vs. G6Drug	-0.4396	-0.4585 to -0.4207	****	<0.0001
BxM	TBZ	G2 vs. G6Uns	0.1848	0.1665 to 0.2031	****	<0.0001
BxM	TBZ	G3 vs. G4	-0.3044	-0.3386 to -0.2702	****	<0.0001
BxM	TBZ	G3 vs. G5	-0.2709	-0.303 to -0.2388	****	<0.0001
BxM	TBZ	G3 vs. G6Drug	-0.5658	-0.5929 to -0.5387	****	<0.0001
BxM	TBZ	G3 vs. G6Uns	0.0586	0.03194 to 0.08526	****	<0.0001
BxM	TBZ	G4 vs. G6Drug	-0.2614	-0.2903 to -0.2325	****	<0.0001
BxM	TBZ	G4 vs. G6Uns	0.363	0.3344 to 0.3916	****	<0.0001
BxM	TBZ	G5 vs. G6Drug	-0.2949	-0.3213 to -0.2685	****	<0.0001
BxM	TBZ	G5 vs. G6Uns	0.3295	0.3035 to 0.3555	****	<0.0001
BxM	TBZ	G6Drug vs. G6Uns	0.6244	0.605 to 0.6438	****	<0.0001

BxM	MPTL	McMaster1931 vs. G1_high	-5.126	-8.498 to -1.755	****	<0.0001
BxM	MPTL	McMaster1931 vs. G2_high	-9.396	-12.74 to -6.05	****	<0.0001
BxM	MPTL	McMaster1931 vs. G3_high	-4.076	-7.725 to -0.4274	*	0.0132
BxM	MPTL	McMaster1931 vs. G4_high	-4.153	-7.837 to -0.4695	*	0.0116
BxM	MPTL	McMaster1931 vs. G5	-4.164	-7.82 to -0.5088	*	0.01
BxM	MPTL	McMaster1931 vs. G6_low	-5.615	-9.264 to -1.966	****	<0.0001
BxM	MPTL	McMaster1931 vs. G6Uns_high	-4.752	-8.197 to -1.307	***	0.0003
BxM	MPTL	Bundarra2014 vs. G1_low	2.514	0.2554 to 4.773	*	0.0139
BxM	MPTL	Bundarra2014 vs. G1_high	-2.61	-4.869 to -0.3514	**	0.0081
BxM	MPTL	Bundarra2014 vs. G2_low	2.515	0.2947 to 4.735	*	0.0109
BxM	MPTL	Bundarra2014 vs. G2_high	-6.88	-9.1 to -4.66	****	<0.0001
BxM	MPTL	Bundarra2014 vs. G6_low	-3.099	-5.754 to -0.4442	**	0.007
BxM	MPTL	Bundarra2014 vs. G6Uns_low	2.514	0.1477 to 4.881	*	0.0252
BxM	MPTL	G1_low vs. G1_high	-5.124	-6.693 to -3.555	****	<0.0001
BxM	MPTL	G1_low vs. G2_high	-9.394	-10.91 to -7.881	****	<0.0001
BxM	MPTL	G1_low vs. G3_high	-4.074	-6.174 to -1.974	****	<0.0001
BxM	MPTL	G1_low vs. G4_high	-4.151	-6.311 to -1.991	****	<0.0001
BxM	MPTL	G1_low vs. G5	-4.162	-6.273 to -2.051	****	<0.0001
BxM	MPTL	G1_low vs. G6_low	-5.613	-7.713 to -3.513	****	<0.0001
BxM	MPTL	G1_low vs. G6Uns_high	-4.75	-6.47 to -3.029	****	<0.0001
BxM	MPTL	G1_high vs. G2_low	5.125	3.612 to 6.638	****	<0.0001
BxM	MPTL	G1_high vs. G2_high	-4.27	-5.783 to -2.757	****	<0.0001
BxM	MPTL	G1_high vs. G3_low	5.123	3.024 to 7.223	****	<0.0001
BxM	MPTL	G1_high vs. G4_low	5.124	2.965 to 7.284	****	<0.0001
BxM	MPTL	G1_high vs. G6Uns_low	5.124	3.404 to 6.845	****	<0.0001
BxM	MPTL	G2_low vs. G2_high	-9.395	-10.85 to -7.94	****	<0.0001
BxM	MPTL	G2_low vs. G3_high	-4.075	-6.133 to -2.017	****	<0.0001
BxM	MPTL	G2_low vs. G4_high	-4.152	-6.272 to -2.032	****	<0.0001
BxM	MPTL	G2_low vs. G5	-4.163	-6.233 to -2.093	****	<0.0001
BxM	MPTL	G2_low vs. G6_low	-5.614	-7.672 to -3.556	****	<0.0001
BxM	MPTL	G2_low vs. G6Uns_high	-4.751	-6.421 to -3.08	****	<0.0001
BxM	MPTL	G2_high vs. G3_low	9.393	7.335 to 11.45	****	<0.0001
BxM	MPTL	G2_high vs. G3_high	5.32	3.262 to 7.378	****	<0.0001
BxM	MPTL	G2_high vs. G4_low	9.394	7.275 to 11.51	****	<0.0001
BxM	MPTL	G2_high vs. G4_high	5.243	3.123 to 7.363	****	<0.0001
BxM	MPTL	G2_high vs. G5	5.232	3.162 to 7.302	****	<0.0001
BxM	MPTL	G2_high vs. G6_low	3.781	1.723 to 5.839	****	<0.0001
BxM	MPTL	G2_high vs. G6Uns_low	9.394	7.724 to 11.06	****	<0.0001
BxM	MPTL	G2_high vs. G6Uns_high	4.645	2.974 to 6.315	****	<0.0001
BxM	MPTL	G3_low vs. G3_high	-4.073	-6.594 to -1.552	****	<0.0001

BxM	MPTL	G3_low vs. G4_high	-4.15	-6.722 to -1.579	****	<0.0001
BxM	MPTL	G3_low vs. G5	-4.161	-6.692 to -1.631	****	<0.0001
BxM	MPTL	G3_low vs. G6_low	-5.612	-8.133 to -3.091	****	<0.0001
BxM	MPTL	G3_low vs. G6Uns_high	-4.749	-6.964 to -2.533	****	<0.0001
BxM	MPTL	G3_high vs. G4_low	4.074	1.503 to 6.645	****	<0.0001
BxM	MPTL	G3_high vs. G6Uns_low	4.074	1.859 to 6.29	****	<0.0001
BxM	MPTL	G4_low vs. G4_high	-4.151	-6.772 to -1.531	****	<0.0001
BxM	MPTL	G4_low vs. G5	-4.162	-6.743 to -1.582	****	<0.0001
BxM	MPTL	G4_low vs. G6_low	-5.613	-8.184 to -3.042	****	<0.0001
BxM	MPTL	G4_low vs. G6Uns_high	-4.75	-7.022 to -2.477	****	<0.0001
BxM	MPTL	G4_high vs. G6Uns_low	4.151	1.879 to 6.424	****	<0.0001
BxM	MPTL	G5 vs. G6Uns_low	4.162	1.936 to 6.389	****	<0.0001
BxM	MPTL	G6_low vs. G6Uns_low	5.613	3.398 to 7.829	****	<0.0001
BxM	MPTL	G6Uns_low vs. G6Uns_high	-4.75	-6.61 to -2.889	****	<0.0001

9.8 SNP allelotyping raw data

This appendix is provided as a separate csv file (BAHE0315_Appendix9.8_SNPrawData.csv) and not included in this document due to its size.

The sequences of primers used for amplification and primer extension and the sets which go together to form each "plex", and allele frequency value associated with drug resistance is sensitive intellectual property which cannot be protected except as a "trade secret".

Appendix 9.8 unavailable in public access of this report due to IP.

9.9 Field evaluation of the genetic drive chip project proposal

Appendix 9.9 unavailable in public access of this report due to IP.