Project No.:	ON-00373
Contract No.:	PO4500007843
AWI Project Manager:	Bridget Peachey
Contractor Name:	The University of Melbourne
Prepared by:	Trent Perry & Philip Batterham
Publication date:	21 November 2018

Genetics of Blowfly Parasitism

Published by Australian Wool Innovation Limited, Level 6, 68 Harrington Street, THE ROCKS, NSW, 2000

 Australian Wool Innovation Limited gratefully acknowledges the funds provided by the Australian government to support the research, development and innovation detailed in this publication.

Table of Contents

Executive Summary

The main goals of this research project were to gain a greater understanding of the blowfly genes involved in targeting and parasitizing sheep. Genes that allow blowflies to find sheep and those essential for larval development on sheep were examined. Bioinformatic analysis of the genes expressed in early development was performed to predict those that might be useful candidates for blowfly control, particularly through vaccine development.

Testing of some of these candidates through gene knockdown to test how vital their protein products were for larval development was conducted. The study has provided a list of candidate genes, prioritised from a combination of biology and bioinformatics. Further work will be required to validate their function and to assess their effectiveness as antigens for vaccine development. Population samples from across the country were also collected and examined for variation in these candidate genes as a low level of variation is an absolute prerequisite for blowfly proteins that could be used as a vaccine. Thus, the population studies provided a valuable sieving process. The desired low level of sequence variation in sequence between populations was found for many of the candidate genes. Other genes displayed high variability, or variation that would render them ineffective as a vaccine antigen. These can now be ruled out as being useful vaccine candidates, saving a lot of needless downstream validation work and highlighting the benefit of having natural population data available.

As part of the analysis of the collection we performed diagnostic assays to identify the proportion of blowflies that have a natural infection of an endosymbiotic bacteria, *Wolbachia*. This bacteria is commonly found in insects and can impact their fertility and behaviour in many different ways. Its presence in blowflies across the country warrants further examination to determine if it plays a role in their parasitism of sheep.

Introduction/Hypotheses

There has been a huge improvement in the technologies available to investigate the biology of organisms and to understand the genes and proteins critical for survival and behaviour. This proposal was designed to use genomics and transcriptomics to investigate blowfly biology at the detection and invasion stages of blowfly parasitism of sheep.

The aims were to identify genes that have potential to facilitate novel control strategies for this pest. The first potential intervention was to investigate the detection of a host, or to understand which genes might help flies find sheep, so this could be interrupted through a bait or trapping strategy. The second area of investigation was to identify genes that were expressed during maggot development prior to establishment of a wound on sheep. The logic behind examining this early stage was that damage from early stage maggots is minor but worsens rapidly. Hence, if an immune response from sheep was to be protective, it would most likely need to be effective at recognising and responding to blowfly larvae even before they wound the sheep. Given that exudates containing antibodies are produced by sheep prior to wounding, this may be possible, provided the right target protein can be found. We believed that identification of those genes expressed early and use of bioinformatics as a filtering process would help to identify a much smaller sub-set of genes encoding useful targets that would expedite and reduce the cost of identifying promising vaccine candidates.

Project Objectives

The project focussed on several objectives to enhance our understanding of the blowfly, improve the quality of genomic resouces available for its study and to identify genes involved in the parasitism of sheep.

The research objectives of this project were to;

1. Improve the *L. cuprina* genome assembly

2. Examine the genes that would play a role in blowfly odour detection

a) Identify the members of the odorant receptor gene family
b) Create and sequence antennal and maxillary palp transcriptomes to detect differentially expressed odorant receptors
c) Create and sequence wing transcriptomes to detect differentially expressed chemosensory receptors

3. Identify genes expressed during early parasitism

a) Conduct a larval implant study with sampling across development and sequence the transcriptomes of these larvae
b) Analyse transcriptomes of larvae to identify genes as potential vaccine candidates
c) Collect and sequence blowfly natural population samples to examine the sequence variation in candidate genes from 3b.
d) Measure the level of *Wolbachia* infections in natural populations
e) Use dsRNA knockdown to examine candidates from 3b.

Success in Achieving Objectives

Objective 1. Improve the *L. cuprina* genome assembly

The project has assembled and annotated a new version of the *L. cuprina* genome (Freeze 2) which has greatly improved the quality of the genome.

Objective 2. Examine the genes that would play a role in blowfly odour detection

The odorant receptor (OR) genes of *L. cuprina* were extracted from the newly assembled genome. The two tissue specific transcriptomes, one for the antennae and maxillary palp and the other for blowfly wings were sequenced and expression differences in OR expression examined. We were able to show that there are no odorant receptor genes that have expression exclusively in either males vs females or in females that are mated or not. We expanded our analysis to include chemosensory receptors and some of these were identified with differences associated with the adult diet provided. The observed impact of diet on odour/taste detection is a particularly interesting result that merits follow up, including whether the larval diet also has an impact on expression of adult chemosensory genes and possibly their host preference. Further work examining a wing transcriptome identified some additional differences in chemosensory genes that may play a role in selection or attraction to host sites.

Objective 3. Identifying genes expressed during early parasitism

A time course study of gene expression during larval development on sheep was conducted and bioinformatic analysis used to filter these to predict suitable vaccine candidates. dsRNA knockdown experiments were used to examine 10 of these genes, with several presenting with phenotypes that would indicate larval development and survival is affected when the level of gene/protein expression is impacted, indicating this is a tool that can contribute to screening for candidates critical for survival of larvae on sheep.

Methodology

Genomic DNA sequencing

DNA extractions were performed using a standard Phenol:Chloroform extraction protocol following incubation with a CTAB/Protease solution. Samples were analysed for DNA concentration on a Qubit analyser and visualised on Agarose gel. Prior to being sent for sequencing, DNA was run on a fragment analyser to measure purity and ensure DNA was not degraded. Sequencing of the male genome for Freeze 2 was performed at DoveTail genomics. The reads were assembled at UoM using customised genomic assembly pipelines and the gene predictions were made using MAKER3.

Sequencing of the pooled DNA extracted from blowflies from natural populations was performed at the Australian Genome Research Facility, Parkville. Reads were trimmed and cleaned, then mapped to the Freeze2 genome. Mpileup files were created and the Varscan program used to determine the base/SNP calls. Candidate genes were extracted and examined for non-synonomous variation.

RNA sequencing

RNA was prepared for transcriptomes using the Promega Reliaprep RNA column kit as per manufacturer's instructions, with samples analysed for concentration on a Qubit analyser and visualised on an agarose gel. A fragment analyser was used to measure the quality of the RNA prior to shipping to Custom Science in Gentegra RNA stabilisation tubes. All RNA sequencing (Antennal/Maxillary palp, Wing and Larval transcriptomes) was performed by Custom Science. RNA reads were quality checked, trimmed and filtered prior to being mapped to the Freeze 2 genome assembly using custom UoM pipelines. The expression counts were compiled together with annotation of their predicted function using KEGG, Swissprot, NCBI and other online databases. Analysis of gene expression and cluster analysis was performed using customised scripts in RStudio.

Gene knockdown

Primers were designed to specifically amplify target candidate genes. RNA was extracted using the Trisure extraction protocol. cDNA was produced using the SSIII reverse transcription kit as per manufacturers protocols for the oligodT method. T7 RNA polymerase binding sites were added to primers that amplified correct fragment sizes and these were then used to amplify a DNA template for RNA production which were purified and quantified using the Qubit analyser. dsRNA was produced using the HiScribe T7 High Yield RNA Synthesis Kit. RNA was purified using LiCl precipitation and quantified using the Qubit analyser followed by agarose gel visualisation for confirmation of product size and quality.

dsRNA was microinjected into embryos and the phenotypes of larvae examined.

Results and Discussion

Objective 1.

A high quality version of the *L. cuprina* genome (Freeze 2) has been assembled and annotated. New long read sequencing using PACBio and Chicago methods has allowed a massive increase in the quality of the genome, compared to the one published in 2015 (Freeze 1). Table 1 shows some revealing statistics comparing Freezes 1 and 2. The average length of DNA scaffolds is far higher for Freeze 2. This means that the full genome sequence is represented in far fewer large scaffolds, rather than many small fragments. More continuity of sequence assists with annotation of complete genes as well as providing information on the gene order. In this context, note the N50 and N90 values which show the number of scaffolds that include 50% and 90% of the complete genome sequence, respectively. The number of predicted genes has decreased to 12,933 genes (Table 2). This is due to the fact that in Freeze 1 many variants of a single gene were counted as different genes. The long read sequence data has allowed us to remove the ambiguities and increase the accuracy of our gene calls. The quality of the Freeze 2 genome is vastly superior when compared to other pest insect genomes such as housefly (N50 = 0.23Mb) and compares favourably with the genome of the model genetic organism, *Drosophila melanogaster*, (N50 = 23.01Mb) in which over 100 years of genetics research and hundreds of millions of dollars have been invested. The final gene set has been submitted, quality checked and accepted by the National Center for Biotechnology Information (USA) which will provide public access to the genome data, following the acceptance of our publication of this new version of the genome that is in preparation. The number of genes without orthologues has also decreased substantially due to the large amount of insect species that have been sequenced over the last 3 years.

Table 1. Key statistics for the *L. cuprina* genome assembly

Features of the Genome assemblies	2015 published genome (458 Mb)	2018 Freeze 2 genome (465 Mb)
N50 length (bp);	744,413bp;	6,922,854bp;
Scaffolds required to reach 50% coverage	**165 Scaffolds**	**18 Scaffolds**
N90 length (bp);	126,471bp;	1,321,550bp;
Scaffolds required to reach 90% coverage	**736 Scaffolds**	**83 Scaffolds**
BUSCO complete genes;	2594;	2704;
BUSCO fragmented genes;	52;	47;
BUSCO missing genes	**153**	**48**

N50 = The scaffold length such that the sum of the lengths of all scaffolds of this size or larger is equal to 50% of the total assembly length.

BUSCO = Benchmarking Universal Single-Copy Orthologues and is a set of genes that is used as a measure of genome quality/completeness.

Table 2. Comparison of gene predictions from 2015 and 2018 genome assemblies

Gene Prediction Comparison	2015 Genome (14,544 genes)	2018 Genome (12,933 genes)
Genes supported by expression data	10,121 genes	10,065 genes
Single-copy orthologues (4 spp.)	4,106 genes	4,425 genes
Single-copy orthologues (1 sp.)	12,160 genes	11,142 genes
Genes unique to the blowfly	2,062 genes	572 genes

Objective 2.

Objective 2a) Identify the members of the odorant receptor gene family

The family of *L. cuprina* odorant receptor (OR) genes has been identified and extracted from the genome.

Table 3. Comparison of odorant receptor genes identified in blowfly genomes.

	L. cuprina 2015 genome	*L. cuprina* Freeze 2	*Calliphora stygia* transcriptome
Number of genes	54	47	50

As shown in Table 3, there were seven fewer ORs annotated in the Freeze 2 of the genome that is likely due to the prior assembly incorrectly annotating alleles as separate genes that have been resolved with the new Freeze 2 genome assembly. The key odorant co-receptor (*Orco*) was identified was used to assist our work on a separate project (ON-00315) in efforts to knock out this gene. The evolutionary relationship of other *L. cuprina* OR genes has also been examined, particularly the expansion of particular receptor clades in blowflies vs *D. melanogaster* (Figure 1). The different colours show *D.melanogaster* (red), *L.cuprina* (brown) and *C.stygia* (blue) OR genes. While there are a number of 1:1 orthologues (genes that are highly similar in protein structure) our analysis has also determined that there appear to be some blowfly specific gene expansions while there are other Drosophila specific OR gene expansions. The biological significance of this is one area of study that could be looked at in further work.

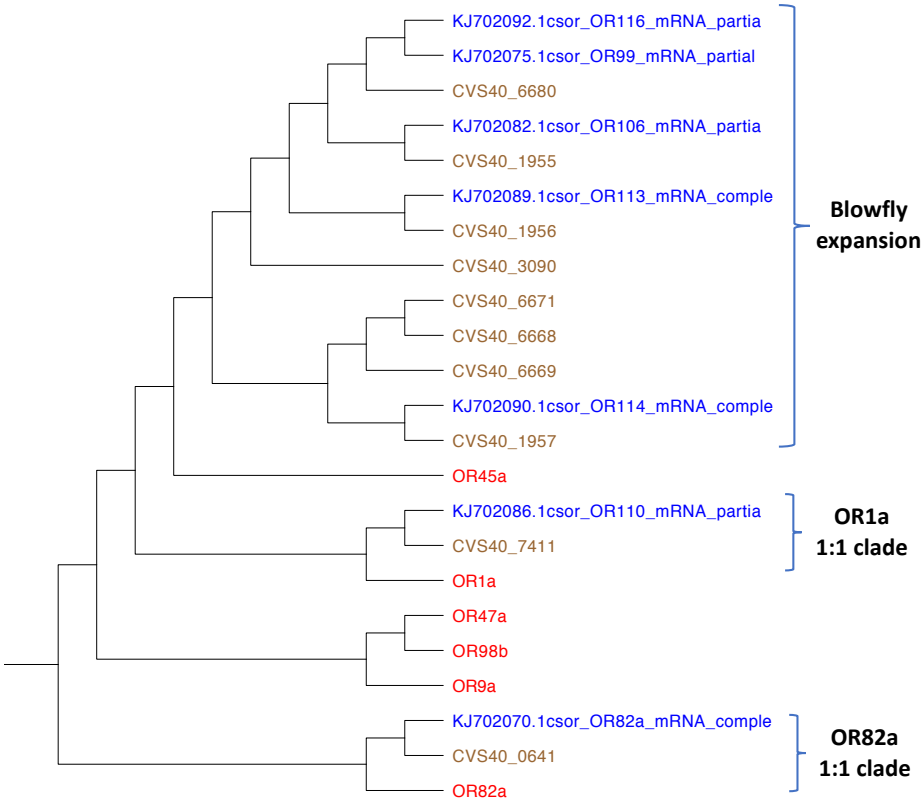

Figure 1. Extract from our analysis of the phylogenetic trees generated from comparison of the OR gene sequences of *D.melanogaster* (red), *L.cuprina* (brown) and *C.stygia* (blue). Some odorant receptors have clear 1:1 relationships (i.e. orthologues) between the three species while for other cases there are either Drosophila or blowfly-specific lineage expansions.

These results will be included in a manuscript on *L. cuprina* ORs, together with data from the olfactory organ transcriptomes for submission early in 2019 (after approval from AWI).

Objective 2b. Dissection, sequencing and analysis of antennal and maxillary palp transcriptomes to look for differences in expression of odorant receptors in adults

A recent report on *Calliphora stygia* had identified female specific OR gene expression. Given the reported prevalence of female blowflies on sheep, we examined the blowfly for differences in expression of OR genes to see if there were sex specific differences in ORs or differences in expression between unmated females and gravid females. OR gene expression profiles from antennae and maxillary palps of blowflies reared under different conditions was performed to identify genes that might be differentially expressed between males, gravid females and females that had not mated. Females that were only reared on agar/sucrose diet were also included to control for expression changes that might be due to protein in the diet which is critical for female reproductive behaviour. Thus, four groups of flies were reared for dissection and the extraction of RNA from the major olfactory organs, the antennae and maxillary palp organs. These were from:-

- Mated females fed on meat (MFM)
- Mated males fed on meat (MMM)
- Unmated females fed on meat (UFM)
- Unmated females reared on agarose (UFA)

The antennae and maxillary palp were removed from 20 adult flies for each of these four groups and RNA extracted from each of three separate replicates. The RNA libraries were prepared and sequenced at Novogene and the data returned to us for analysis. The data were filtered, and sequencing adaptors trimmed before the reads were mapped to the genome to create a table of gene counts. Given that there are several chemosensory proteins known, the gustatory (taste) and ionotropic receptors (i.e. bitterness, salt, ammonia) were analysed in addition to the ORs. In all, a total of 99 genes were analysed. 9 of these did not have any detectable expression in the tissues examined and thus were excluded from further analysis. Sequencing data for the remaining 90 chemosensory proteins analysed are presented below. Naming of the genes has followed convention from the literature with the names reflecting the closest *D. melanogaster* OR gene orthologue.

Looking at the biological coefficient of variation plot (Figure 2) we can see how much of the variation in the samples is explained by particular factors. We had expected to see gender based differences but did not find them. The real differences that we see can be attributed to differences in adult diet, particularly whether they had access to meat. Adult flies fed meat, whether male or female, mated or unmated, are more similar than the females only reared on agar. Of themselves these data would prompt us to ask if flies are primed through their diet to find sheep and if there are additional mechanisms that contribute to the reported switch that gravid females undergo to begin host-seeking behaviour.

Representation of relatedness of samples based on variation in gene expression

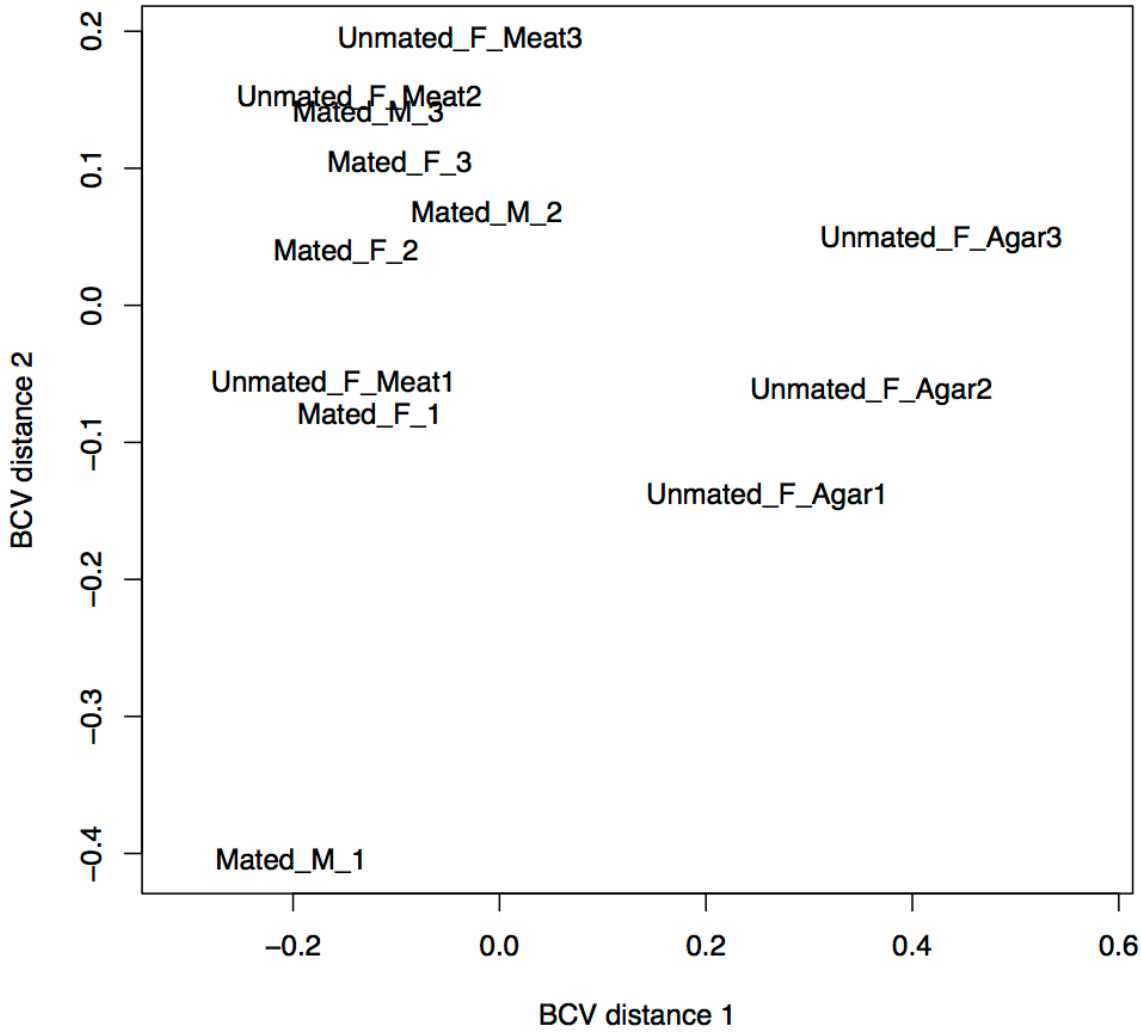

Figure 2. The biological differences between samples are represented on the above plot. The further apart samples are, the more different their expression profile is. We can see that the most distinct cluster of samples are the unmated females on Agar media, all clustering towards the right of the plot. This highlights that the difference in the chemosensory genes appears to be most influenced by the diet of the flies rather than sex or if they have been mated or not.

Heat map of the antennal/maxillary palp sample gene expression

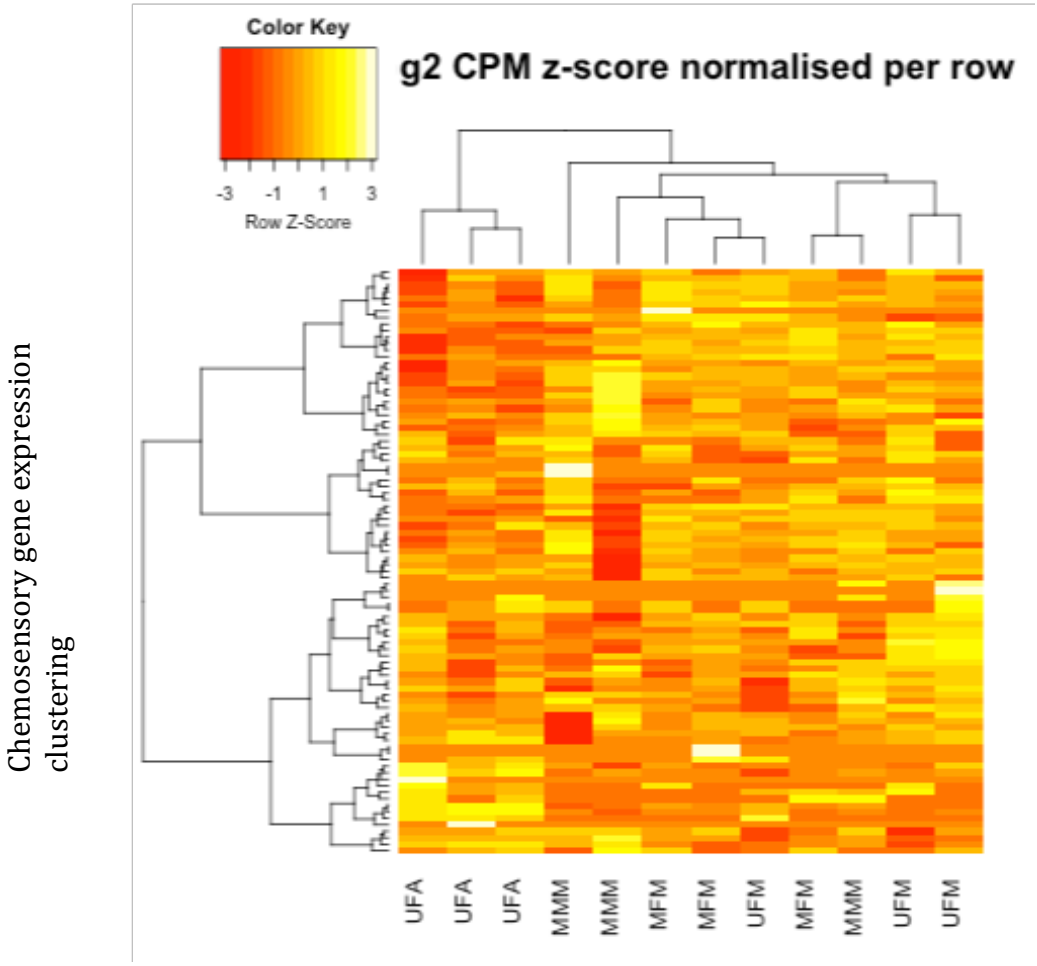

Figure 3. Heat map of the antennal/maxillary palp sample gene expression in counts per million (CPM) (X-axis), relative to the 90 chemosensory genes analysed (Y-axis). The different conditions (3 separate replicates) are clustered horizontally according to similarity in expression profiles and the genes are clustered down the vertical axis according to their similarity in expression across the samples.

Figure 3 presents a cluster analysis of the levels of gene expression that was performed on all the samples for the 90 chemosensory genes with expression counts. The levels of expression are represented as a heat map of red (low expression) to white (high expression). As with the BCV plot in Figure 2, we observe a clustering of the agar reared unmated females (UFA), but the samples from meat fed flies are not distinct at the overall chemosensory gene expression level as seen by the mixed order along the X-axis of MMM, MFM, and UFM samples.

Grouping of similar gene expression patterns based on hierarchical clustering

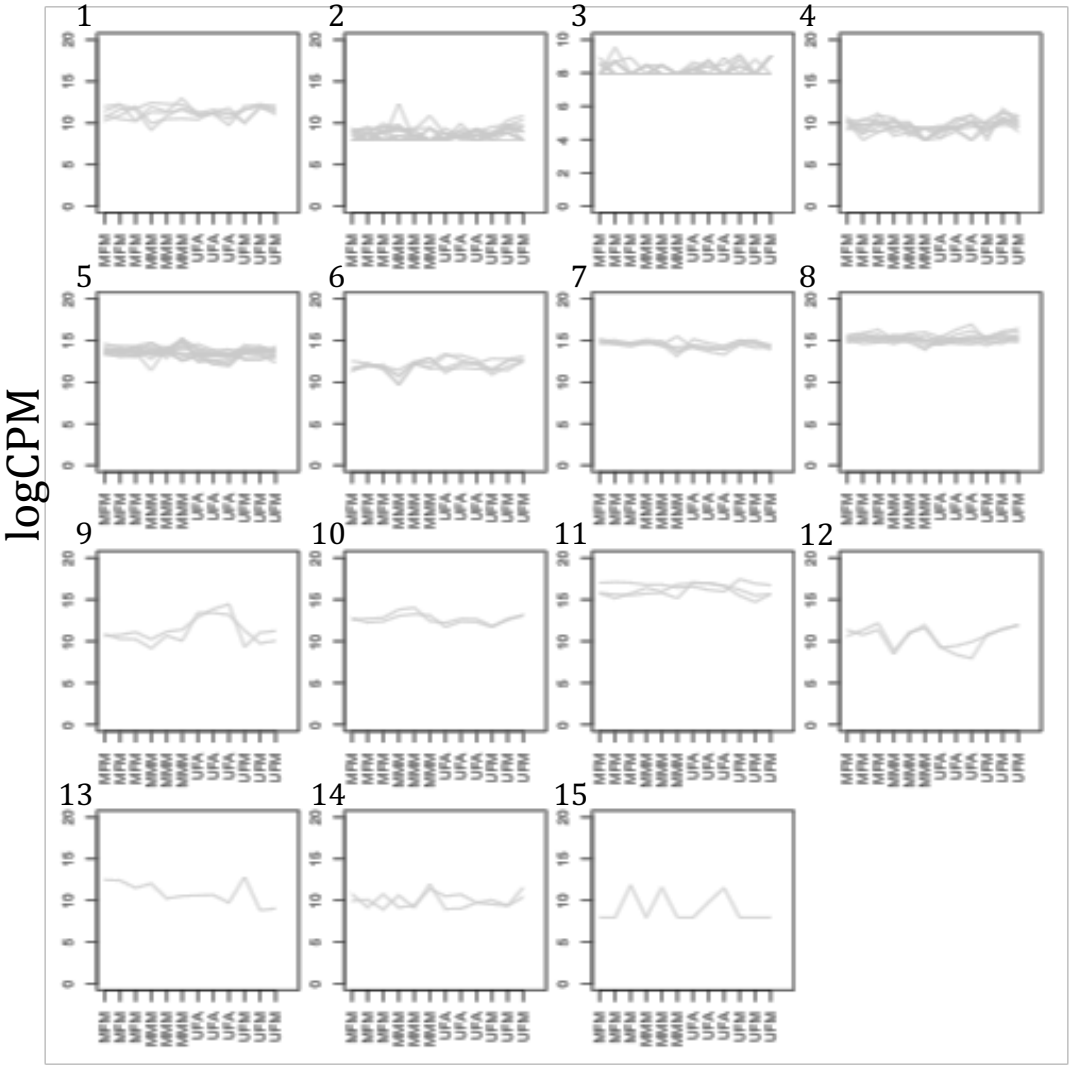

Figure 4. Plots of individual gene expression from the chemosensory gene set based on similarity of expression pattern between all samples. 15 clusters were assigned. On the X-axis are the samples in replicates of 3 for MFM, MMM, UFA and UFM. The Y-axis represents LogCPM. Each line is an individual gene.

The results from the clustering analysis can also be broken down into graphs of genes with similar patterns of expression. Figure 4 contains these plots of the chemosensory gene expression counts divided among 15 clusters (sets of genes showing similar expression patterns). The genes found to have expression patterns significantly different between various conditions are reported below.

The gene that was found to be most different in expression levels in mated females fed on meat relative to most replicates of other samples is the putative gustatory receptor 8a, seen in cluster plot 13. In *D. melanogaster* the orthologue is required for the aversive response to the compound L-canavanine. Gr8a expression is generally higher in mated females fed meat relative to other conditions. As yet there are no data suggesting a response to other

compounds. Thus, there could be a role for this subunit in detecting a particular odour or compound. It would be worth investigating whether the response to L-canavanine in *L. cuprina* is one of attraction or aversion, as this could link a behavioural response to detection of the compound. Two genes have higher expression in both male and female flies fed meat. Both of these genes are odorant receptors, Or7a-like and Or59a-like. Given these are all meat fed, this is a significant finding, however as yet we do not know the specific odours that these receptors respond to, this would require further research. Males and females were not found to express different odorant receptors in the antennal/maxillary palp libraries, but the gustatory receptor, Gr58a appears more highly expressed in unmated or mated females fed meat than in mated males fed meat or unmated females on agar.

The results from analysis on only the odorant receptors indicate that there does not appear to be a difference between expression of the odorant receptors between the mated males and females fed on meat. There are, however, some significant differences in expression that have been detected in expression levels of other chemosensory genes. The gene most highly expressed in a comparison between meat fed females compared to meat fed males and unmated females is a gustatory protein. The closest match to this predicted gene in insect databases is a gustatory receptor, Gr58a-like. In the literature, there are no reports of ligands that this receptor responds to.

There are two genes more highly expressed in the UFA samples, relative to all other conditions (Figure 4, cluster 9). One is the putative gustatory receptor 64f that has been reported to mediate responses to various sugars while the other is putative odorant receptor 67d. The orthologues of Or67d have been shown to be involved in promoting female receptivity in *D. melanogaster*. The fact that females fed meat but not mated, do not also have high levels of this gene could indicate it is downregulated when protein sources are available. Analysis of mating behaviour and response to different odours for this odorant receptor might be able to address this hypothesis. Table 4 (*below*) summarises the differences in expression observed in the chemosensory gene analysis of the transcriptomes.

Follow up work on these genes could include further sampling to test a greater range of conditions helping to validate these findings and work on identifying the odorants and the nature of the responses blowflies have when they are detected (ie. behaviours such as attraction).

Table 4. Summary of differences in chemosensory gene expression

Gene	Expression pattern	Phenotype	Cluster
putative gustatory receptor 8a	Higher in MFM vs all	Gr8a is required for response to the aversive compound L-canavanine but not other compounds	13
putative gustatory receptor 64f	Higher in UFA vs all	responses to various sugars	9
putative odorant receptor 67d	Higher in UFA vs all	promotes female receptivity and inhibits male-male courtship	9
putative gustatory receptor 58a	Higher in MFM & UFM vs UFA and MMM	-	4
putative odorant receptor 7a	Lower in UFA vs all	-	12
Odorant receptor 59a	Lower in UFA vs all	-	12

Objective 2c. Create and sequence wing transcriptomes to detect differentially expressed chemosensory receptors

Given the lack of difference between the male and female samples from the antennal/maxillary palp transcriptome analysis, we examined the chemosensory genes expressed in the wings of the blowfly. These are part of insect detection and guidance systems and have been shown to play a role in aggregation and also egg laying. Research from *D. melanogaster* indicated that gustatory sensilla (taste perception organs) are located on the anterior wing margin. Further work identified the presence of gustatory receptors (GR) in wings of two other species, the honeybee (*Apis mellifera*) and aphid (*Acyrtosiphon psium*). This suggested that there could be a role of these in taste perception and a role of these receptors in insect wings in the detection of sites for mating, feeding and laying of eggs. These studies identified these sensilla as responding to sweet and bitter molecules. When the insects were exposed to pulverised sugar droplets that normal behaviour was aggregation and egg laying. This behaviour was reduced when wing taste sensilla function was impaired.

Following a similar treatment regime to the antennal/maxillary palp transcriptomes, wings were prepared, RNA extracted and then sequenced. The samples were again prepared from:-

- Mated females fed on meat (MFM)
- Mated males fed on meat (MMM)
- Unmated females fed on meat (UFM)
- Unmated females reared on agarose (UFA)

Levels of RNA transcript expression in the wings are low, however sufficient RNA was isolated from 50 dissected wings per sample to allow sequencing. Analysis of all samples detected expression of 4 GR genes in the wings, GR94a and GR98a that are involved with sweet taste receptor activity while Gr59f and Gr32a are associated with bitter taste receptor activity. From the different samples, the GR59f bitter taste receptor was most highly expressed. For both bitter taste receptor genes, there is a higher level of expression in the UFA wing samples compared to other conditions (GR59f 1.5-fold, GR32a 2 to 4-fold).

There was not a significant overlap in the expression differences identified between the two tissue specific transcriptomes. We highlighted several genes that might be worth further investigation, however unlike other species, it does not appear that there is sex-specific expression of ORs as was found in a carrion-dependent blowfly, *C. stygia*. This is despite us identifying the orthologues in *L. cuprina* (CVS40_8174 and CVS40_3333) for the two sex specific OR genes (OR118 and OR119, respectively) described in that study.

Overall our findings have indicated that there is a lot of fundamental research that needs to be conducted to better understand olfaction and its link to behaviour in the sheep blowfly.

Objective 3.

Objective 3a. Conduct a larval implant study with sampling across development and sequence the transcriptomes of these larvae

Approval was received for our implant study from the FVAS Ethics Committee (Ethics ID: 1613853.1). The study involved implantation of blowfly embryos onto three sheep at six independent sites. We also used the placement of eggs on meat and the cotton plugs for comparison. Developing animals were harvested from these implant sites over a 22 hour period, leading to samples of embryos and larvae that were frozen in liquid nitrogen. RNA was extracted from samples, quality checked, shipped and sequenced by Custom Science. The reads were mapped to the Freeze 2 genome and transcript counts calculated prior to analysis.

Objective 3b. Analyse transcriptomes of larvae to identify genes as potential vaccine candidates

We examined the expression levels of genes across the sampling period to identify those that were consistently highly expressed during this early pre-parasitic phase of development. The genes were first filtered to identify the subsets of encoded proteins that were (a) excretory/secretory proteins, (b) have predicted transmembrane domains or (c) orphans (not previously identified in other species). The proteins that have transmembrane domains and the predicted excretory/secretory proteins are most likely to have regions that would be accessible for recognition by the host immune system. The orphan genes were also included in the analysis as their function cannot be predicted. They may have novel functions that allow the blowfly to parasitise sheep and thus they could be useful specific targets for the blowfly.

Specifically, we looked for genes expressed very early in larvae, as soon as the larvae hatched from embryos. Considering the ultimate goal of developing a vaccine, it would be most effective if the vaccine disrupted the function of a protein required for the viability of the fly as early as possible in development with the target being present long enough for a response to be mounted by the sheep. The genes were also filtered to identify ones that remained highly expressed in the later stage larvae that were beginning to feed at the skin and which also maintained high levels of expression when reared exclusively on meat. For the immune system to build up antigens and respond, the protein needs to be present early to stimulate the immune response, but also remain expressed and be critical for survival and growth so any immune response will detrimentally impact developing larvae.

The genes matching our filtering criteria described above with the highest, most consistent expression were then examined for known orthologues. Of particular interest were genes where a loss of function in their orthologues in another species had been shown to reduce viability or lead to lethality. Most of the available data are from *D. melanogaster* where a large number of genes have been analysed for loss of function phenotypes. Of our candidate proteins, seven genes are reported to cause a lethal phenotype when knocked down in Drosophila. As we will discuss later, we have found that these data do not always match up with our gene knockdown results in the blowfly itself. This indicates that while Drosophila might be a useful guide, validation in the blowfly is necessary to confirm if a particular gene is required for blowfly survival.

Objective 3c. Natural population sampling

Another level of filtering of the candidate list was to compare the levels of protein sequence variation in the set of candidate genes. Proteins that carry out vital functions are evolutionarily constrained and would be expected to show low levels of variability. To do this, pooled samples from natural populations collected in VIC, NSW, QLD and WA (Figure 5). *L. cuprina* adults that were sampled from 11 sites are listed (Table 6). These pools of flies were sequenced and the levels of variation in our set of candidate genes analysed.

The candidate proteins had "% difference" in protein variation ranging from 0% to 7.39% in blowflies from populations from sites around Australia. 11 of the candidates had less than 1% difference in protein sequence. In cases of high variation, it is important to consider where on the protein the variation is occurring. If exposed outer surfaces of the proteins or extracellular domains are invariable, these would not be expected to impede the use of a candidate protein as an antigen that would be effective across the country. Thus, specific antigenic fragments would need to be designed to these highly conserved regions for these proteins. More extreme variants, such as those with premature stop codons in the protein, were identified and were removed from our list. These alleles suggest these proteins were not required for flies to survive. It also demonstrated that there are already flies present in the field lacking the target protein. A vaccine based on such a protein would be ineffective.

Table 6. Pooled samples and source locations.

NSWpop	VICpop_1	VICpop_2	QLDpop	WApop
Armidale	Stradbroke	Penhurst	Dutton Park	Mt Barker
Gunningbland	Stratford		Robertson	
Temora	Traralgon		Woolongabba	

Figure 5. The five sites of collections for the sequenced blowfly populations that were used to examine candidate gene variation levels. Sites in each of the pooled collections are listed in Table 6.

Objective 3d. Level of *Wolbachia* infection in natural populations

From the 2017/2018 collection, 154 samples were analysed with a diagnostic PCR of two *Wolbachia* specific amplicons. Of the samples in the collection, 81% had *Wolbachia* present. All states sampled had blowflies carrying a *Wolbachia* infection (Table 7). There is a need to expand the sample numbers and collection sites, particularly in Tasmania and Western Australia where flies were sampled at only 2 sites.

Table 7. Infections detected in *L. cuprina* samples from by state.

	QLD	VIC	NSW	TAS	WA
Total samples	18	80	26	7	23
Infected	9	63	24	7	22
% infection	50%	79%	92%	100%	96%

Sequencing was performed on two different sized PCR amplicons which showed that there is more than one *Wolbachia* strain present in Australian blowfly populations. There is an increasing interest in *Wolbachia* and endosymbionts and their interaction with their insect host biology. Three important conclusions can be reached on the basis of these data:-

1. *L. cuprina* have the capacity to harbour Wolbachia, meaning that the potential to use an appropriate strain of Wolbachia as a control agent exists.
2. That *Wolbachia* is already present in such a high percentage of blowflies, means that *Wolbachia* should be viewed as a natural control solution.
3. While it has not been examined formally, the endemic Wolbachia do not seem to be exhibiting a significant level of control on blowfly populations. Thus, if *Wolbachia* was to be used for control, an effective strain would need to be identified and introduced into the field.

Objective 3e. Examining identified blowfly candidates through gene knockdown

The dsRNA experiments were used to gain an understanding of how to best perform and measure the effects of a reduction in gene function on larval development for validating candidate genes. We used three measures to understand how the survival of the larvae was impacted. The first was the level of hatching, the second the developmental stage of arrest for those that did not hatch and the final measure was how well larvae survived over two days after injection. The injection survival rates are variable between different egg batches and so matched controls were used where possible. Several genes displayed lower levels of hatching which could indicate the dsRNA is severely impacting their development. We also determined that the concentration of dsRNA also affects the survival levels. Hence, for examining candidates for their importance to larval survival it would be recommended that a starting concentration of 500ng/ul or less should be used. At concentrations approaching 1ug/ul there may be detrimental effects that make it difficult to differentiate between dsRNA toxicity and the effect of the gene knockdown on larval development and survival.

Four of the candidates tested did not appear to have significant differences in their survival relative to their controls. One candidate had both low levels of development (11% vs 20% control) and low hatching (16.8% vs 32.7% control). Another had poor levels of hatching (3.43%), however this may have been due to the high dsRNA concentration (1ug/ul) while two of the candidates tested were more promising, with low dsRNA concentrations still leading to a reduced hatching rate 3.4% (vs 27.2% control) and 6.64% (vs 30.11% control) respectively.

One of the candidates was tested three times at two different concentrations and in each experiment was found to cause a high level of mortality, i.e. larvae that did hatch. Effects were noted on larval feeding behaviour, with larvae wandering away from the meat on the first day after hatching and almost all of the injected larvae that hatched dying within 48 hours. A second candidate displaying the aberrant wandering behaviour and dying larvae within 48 hours of hatching was also repeated with similar results.

One limitation of the gene knockdown technique is that we cannot determine whether the mortality is due to it knocking down expression throughout the embryo. This is an example of where the capacity to perform tissue specific expression of RNA silencing constructs would be beneficial. It would allow the impact of gene knockdown in specific tissues to be individually assessed. We also note that one of the candidate genes is not found to be lethal when knocked down in *D. melanogaster*, but was in our experiments on the blowfly. This

means that there might be differences in gene expression pattern, or in the role this protein plays in *L. cuprina* that results in a lethal phenotype. This highlights the value in being able to examine candidate genes in the blowfly itself, ensuring results are as biologically accurate as possible.

Overall the dsRNA results have demonstrated the importance of several candidate genes for larval development and survival. This is a good indication that our bioinformatic approach to candidate filtering and identification has been successful in detecting genes required for larval development and survivial. The screen we have developed to examine the effects of gene knockdown appears to be an extremely useful tool for early validation of candidates and also provides the capacity to titrate the level of dsRNA to examine how different levels of knockdown impact the larvae. This could be a useful guide as to how sensitive to an immune response the larvae might be for a particular candidate. Further work improving blowfly genetic manipulation technologies to enhance the level of throughput and also to allow more precise manipulations would be one of the areas of research where further work would assist with this. A comprehensive analysis of protein expression and secretion levels during parasitism in a time course study similar to the one performed for the transcriptome study would help to prioritise the candidate list that we have and could also extend the list through identification of new candidates performing unique functions in the blowfly.

Impact on Wool Industry – Now & In Five Years' Time

The publication of the original genome sequence has already stimulated renewed interest in blowfly control research. The improved genome and addition of gene expression data will stimulate more research ideas and provide resources that can be used and analysed by other researchers, bringing new ideas to this area. The genome has been a stepping stone for other technical work such as establishing CRISPR technologies and will underpin the adoption and development of further genetic technologies that will allow more targeted analysis of genes of interest. Increasing the quality of the genome will enhance the capacity of researchers to perform these types of study. Our approach to identifying genes that are critical for parasitism and larval survival has generated a large amount of data. We hope that both our approach and our data will help support future projects attempting to design novel blowfly control strategies. The discovery of Wolbachia in blowfly populations opens the door for this bacterium to be investigated as a possible biological control agent.

Commercialisation Report

As reported previously, we have presented the current approach and also future ideas on identifying vaccine candidates with two companies. In both instances it was clear that the project is at too early a stage to attract a commercial partner. There is a need for further biological evidence of a particular gene target being useful as a vaccine candidate. It was not necessarily a requirement that the antigen has already undergone sheep trials. However, this would make the most compelling case for significant commercial investment. The candidate list that is being assembled and analysed in this project is thus at too early a point in the validation process for commercialisation. It does have the potential to be commercially valuable with some further analysis and biological validation that would refine the list to a small number of key candidate genes.

Conclusions and Recommendations

The latest assembly of the *L. cuprina* genome is of high enough quality that, in the near term, deeper sequencing would only provide marginal improvements. The quality of the genome could be enhanced at a reasonable cost once the price of long read sequencing falls further. The quality of the genome will also improve as more research is performed on particular genes and regions of it. This in turn will increase the value of the genome as a resource for researchers. At this stage there would be more benefit gained from work to assign the DNA sequence scaffolds to chromosomes, creating a high resolution genetic map. Such a map would allow the rapid identification of genes of interest, e.g. those conferring insecticide resistance. We are proposing to use and build on the genome resource by examining the blowfly population structure across Australia. Combining this with our transcriptome data will help refine the vaccine candidate list. Further, it will provide a fundamental understanding of the geographical size of blowfly populations and the level of movement between them. For woolgrowers it will answer a basic question of interest – where do the flies come from? For researchers, it will provide essential baseline information required before the field implementation of any control strategy.

This project has shown that there are big differences between *L. cuprina* and other insects in the complement of olfactory genes present and in their expression. The lack of a sex specific expression pattern for any of the olfactory receptors does not mean that there are no olfactory cues that lead to blowfly strike. The olfactory responses could mediate different responses in males and females. This should be examined, as should the extent to which olfaction vs gustatory or other cues enable blowflies use to strike sheep. Work on this would begin with a fly carrying a knockout of the *Orco* gene, work that is being pursued as part of a separate project (ON-00570). The association between OR gene expression and diet needs to be examined, to see if diet impacts the extent to which blowflies are attracted to sheep.

Our implant experiment analysing the gene expression profile of larvae on sheep has led to a large amount of data on genes that are likely to be important for larval development and that could be used as targets for a vaccine. The next phase of work would be to identify those proteins that are excreted by larvae at particular times during development, differentiating between those that are accessible to the sheep immune system and those that are not. Our gene knockdown work to begin examining the candidates on our list demonstrated that we have found a significant number that impact early larval development and survival. Thus, further work to refine and validate the candidate list would be of value. Understanding the importance of, and characterising the role of, these candidates in larval development may also lead to identification of useful targets for other control strategies e.g. chemical or genetic control.

List of Abbreviations and/or Glossary

BUSCO = Benchmarking Universal Single-Copy Orthologues
RNA - Ribonucleic acid
DNA - Deoxyribonucleic acid
gDNA - Genomic DNA
cDNA - complementary DNA
dsRNA – double stranded RNA
PCR - polymerase chain reaction

Appendices

Appendix 1 – List of Milestones & Dates submitted

Milestone	Reporting	Due	Submitted
4003043-0010	Contract Execution	15.06.2016	14.06.2016
4003043-0020	Progress Report	15.11.2016	14.11.2016
4003043-0030	Progress Report	31.03.2017	03.04.2017
4003043-0040	Progress Report	15.07.2017	28.07.2017
4003043-0050	Progress Report	30.04.2018	05.05.2018
4003043-0060	Progress Report	01.07.2018	13.09.2018
4003043-0070	Final Report	15.08.2018	12.11.2018

Appendix 2 – Any Project Intellectual Property

Vaccine candidate gene analysis and validation

1. Individuals who contributed to the invention:

Full name	Contact information	Department	Employee or Contractor?
Trent Perry	trentp@unimelb.edu.au	School of BioSciences	Employee
Philip Batterham	p.batterham@unimelb.edu.au	School of BioSciences	Employee
Vern Bowles	vmb@unimelb.edu.au	Faculty of Veterinary and Agricultural Sciences (FVAS)	Employee
Clare Anstead	clare.anstead@unimelb.edu.au	FVAS	Employee
Robin Gasser	robinbg@unimelb.edu.au	FVAS	Employee
Ross Hall	rossh@unimelb.edu.au	FVAS	Employee
Tinna Yang	ytyang@unimelb.edu.au	School of BioSciences	Employee
Natalia Hernandes	natalia.hernandes@unimelb.edu.au	School of BioSciences	Employee

2. Title of invention:

Vaccine candidate gene analysis and validation

3. Date of first conception (15/06/2015)

Proposal start date describes the initial conception of the process to identify and validate the candidate genes

4. Has the invention been described in specific detail or in a general way in the publication?

Has the invention been disclosed, sold or offered for sale to anyone? If so, please describe to whom and under what conditions (e.g. confidentiality agreement, beta test agreement) and provide relevant dates. **This is extremely important.**

Type of disclosure	**To whom**	**Conditions**	**Date (dd/mm/yyyy)**
Presentation of approach (no candidate gene identities revealed)	Global Vetpharma company	CDA	22/06/2018
Presentation of approach (no candidate gene identities revealed)	CSIRO/MLA	CDA	16/07/2018

List any websites, publications, patents, products, services, etc (i.e. prior art) that you are aware of that are similar to or discusses the subject matter of the invention.

The reverse vaccinology approach has been described in the literature, however the data we have is unique and has not been published.

5. Why do you think this invention is strategically important to the Organisation?

This data would have universal application to other projects investigating gene expression in early development of blowfly larvae. It also has demonstrated an approach that can be scaled up to investigate a large number of genes for their effects on larval survival which could identify genes that could be targeted in new control strategies, not only vaccines, but also as potentially chemical or genetic methods.

6. Are there any Organisation products/services/projects that utilise or may utilise the invention?

There is the opportunity to use the data from this analysis to undertake further gene candidate analysis or to support other projects investigating *L. cuprina* vaccine targets.

7. Additional information: N/A

Appendix 3 – Storage of Primary Research Data (Paper based and Electronic)

The research data or this project has been documented in Laboratory notebooks which are retained on site.

An electronic copy of data, (sequencing, images, etc) and reports has been backed up to the University central drive for secure storage.

Appendix 4 – Animal Ethics Approvals (if any)

Ethics approval was received for the Sheep experiments that involved larval implantation. The Ethics approval number is Ethics ID 1613853.1