Department of Interdisciplinary Life Sciences
University of Veterinary Medicine Vienna

Research Institute of Wildlife Ecology
(Head: Univ.Prof. Dr. rer.nat. Claudia Bieber)

Genomic and Transcriptomic Analyses of *Wohlfahrtia magnifica*, a Myiasis-causing Flesh Fly, Reveal Key Targets for Potential Control Programs

PhD thesis submitted for the fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (PhD)

University of Veterinary Medicine Vienna

submitted by
Zhipeng Jia

Vienna, January 2024

PhD Committee

1st Supervisor

Priv.-Doz. Dr.med.vet. Pamela A. Burger
Research Institute of Wildlife Ecology
Department of Interdisciplinary Life Sciences
University of Veterinary Medicine, Vienna
Vienna, Austria
pamela.burger@vetmeduni.ac.at

2nd Supervisor

Priv.-Doz. Dr.rer.nat. Claus Vogl
Institute of Animal Breeding and Genetics
Department of Biomedical Sciences
University of Veterinary Medicine, Vienna
Vienna, Austria
claus.vogl@vetmeduni.ac.at

3rd Supervisor

Prof. Dr. Surong Hasi
College of Veterinary Medicine
Inner Mongolia Agricultural University
Hohhot, China
surong@imau.edu.cn

List of Publications

Article 1: Jia, Z., Hasi, S., Vogl, C., & Burger, P. A. (2022). Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates. *Molecular Ecology Resources*, 22(7), 2744-2757. https://doi.org/10.1111/1755-0998.13654

Impact factor: 7.7

Article 2: Jia, Z., Hasi, S., Zhan, D., Vogl, C., & Burger, P. A. (2024). Transcriptomic profiling of different developmental stages reveals parasitic strategies of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *BMC Genomics*, 25, 111. https://doi.org/10.1186/s12864-023-09949-3

Impact factor: 4.4

Article 3: Jia, Z., Hasi, S., Zhan, D., Hou, B., Vogl, C., & Burger, P. A. (2023). Genome and transcriptome analyses facilitate genetic control of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *Insects*, 14(7), 620. https://doi.org/10.3390/insects14070620

Impact factor: 3.0

TABLE OF CONTENTS

1. ACKNOWLEDGEMENTS 5

2. AUTHOR CONTRIBUTIONS 6

3. DECLARATION 7

4. SUMMARY 8

5. GENERAL INTRODUCTION 11

6. PUBLICATIONS 25

6.1 Article 1 25

6.2 Article 2 54

6.3 Article 3 65

7. DISCUSSION & CONCLUSION 98

REFERENCES 105

1. ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to all those who have supported and guided me throughout my PhD journey. First and foremost, I would like to express my deep gratitude to my first thesis supervisor, Pamela Burger, for providing me with the opportunity to embark on this challenging yet rewarding journey under her supervision. Her patience, support, encouragement, and dedication to academic excellence have played a fundamental role in influencing this research and my development as a scholar. Without her, this would not have been achievable. I am also grateful to the members of my thesis committee, Claus Vogl and Surong Hasi, for their expert guidance and constructive feedback, which have significantly enriched the quality and depth of this research. My heartfelt gratitude goes to my family, who have been a wellspring of unwavering support, encouragement, and love. Their belief in my capabilities has provided the inspiration and fortitude needed to navigate the challenging path of PhD research. I am deeply appreciative of my friends and colleagues who have contributed in various ways to this academic endeavor. Lastly, this research would not have been feasible without the academic resources, facilities, and scholarly environment provided by the University of Veterinary Medicine, Vienna. I extend my appreciation to the institution, faculty, and staff, particularly within the Research Institute of Wildlife Ecology, for fostering an atmosphere conducive to academic growth. 谢谢你们的支持。

2. AUTHOR CONTRIBUTIONS

Article 1: Jia, Z., Hasi, S., Vogl, C., & Burger, P. A. (2022). Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates. *Molecular Ecology Resources*, 22(7), 2744-2757. https://doi.org/10.1111/1755-0998.13654

Z.J. and P.A.B. conceived the project and received funding. Z.J. performed the genome annotation and comparative genomic analysis and wrote the first draft of the manuscript. S.H. and C.V. contributed new reagents, samples or analytical support. P.A.B., S.H. and C.V. supervised the project. All authors provided valuable advice, reviewed and approved the final manuscript.

Article 2: Jia, Z., Hasi, S., Zhan, D., Vogl, C., & Burger, P. A. (2024). Transcriptomic profiling of different developmental stages reveals parasitic strategies of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *BMC Genomics*, 25, 111. https://doi.org/10.1186/s12864-023-09949-3

Z.J. and P.A.B. conceived and designed this project. S.H. and D.Z. collected the samples. Z.J. performed the data analysis. Z.J. wrote the first draft of the manuscript. P.A.B., C.V. and S.H. revised the manuscript. P.A.B., C.V. and S.H. supervised the project. All authors read and approved the final manuscript.

Article 3: Jia, Z., Hasi, S., Zhan, D., Hou, B., Vogl, C., & Burger, P. A. (2023). Genome and transcriptome analyses facilitate genetic control of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *Insects*, 14(7), 620. https://doi.org/10.3390/insects14070620

Conceptualization, Z.J. and P.A.B.; methodology, Z.J. and P.A.B.; validation, Z.J. and B.H.; data curation, Z.J.; resources, S.H. and D.Z.; writing—original draft preparation, Z.J. and P.A.B.; writing-revision and editing, all authors; supervision, P.A.B., C.V. and S.H.; funding acquisition, Z.J., P.A.B. and S.H. All authors have read and agreed to the published version of the manuscript.

3. DECLARATION

I hereby confirm that I have followed the rules of scientific good practice in all respects.

Zhipeng Jia

January 2024

4. SUMMARY

Background

Wohlfahrtia magnifica, a flesh fly, is one of the most important obligatory traumatic myiasis-causing flies, affecting a range of mammals in several European, African, and Asian countries. The infestation of *W. magnifica* can lead to significant health and welfare problems and substantial economic losses in the field of animal husbandry. However, research on *W. magnifica* is quite limited, especially at the molecular level.

Aims of the thesis

In light of this limited information, the aims of this thesis revolved around filling the existing knowledge gaps in *W. magnifica*. First, I aimed to sequence, assemble, and annotate the genome of *W. magnifica*. This new genome resource provides an essential starting point for many future fundamental and applied research areas. Second, I aimed to investigate the complex dynamics of gene expression throughout the life cycle of *W. magnifica* to understand how parasitic larvae regulate specific gene expressions to effectively parasitize their hosts. This part of my research will guide the selection of potential targets for vaccines or insecticides aimed at disrupting the establishment of the larvae on or in hosts. Third, I aimed to identify and characterize the critical components for the development of genetically modified strains in *W. magnifica*. This represents a significant step in the future establishment of genetic control programs designed to address the infestation of *W. magnifica*.

Results

In our first article, we adopted the strategy of low DNA input library preparation and successfully sequenced the genome of *W. magnifica* using a single adult specimen as the genetic source. The assembled genome spans 753.99 Mb, with a N50 length of 5.00 Mb, and contains 59.71% repeat elements. The overall RNA-seq alignment rate reached 93.62%, and 98.8% of complete BUSCOs (Benchmarking Universal Single-Copy Orthologs) were identified, indicating the completeness and high quality of the genome. The genome annotation

process predicted 16,718 genes and 20,017 mRNA sequences in the genome, with 64.98% of genes functionally annotated using the UniProt/Swiss-Prot database. The phylogenetic analysis demonstrated that *W. magnifica* has the closest relationship to *Sarcophaga bullata*, another member of the Sarcophagidae family, followed by *Lucilia cuprina*, another myiasis-causing fly. The analysis of gene family expansion revealed that the function of the expanded gene families is associated with immunity, insecticide resistance, responses to heat stress, and cuticle development. Furthermore, using the comparative method 45 positively selected genes were identified exhibiting diverse functions.

In the second article, we identified a set of 2049 excretory/secretory (ES) proteins in *W. magnifica*, which play critical roles in parasite-host interactions. Functional annotation indicated that these ES proteins are primarily associated with processes such as cuticle development, peptidase activity, immune responses, and metabolic activities. With the larval samples collected from wounds of the infested host species, Bactrian camel, and the following pupal and adult stages, the analysis of gene expression indicated that the functions of upregulated genes were distributed in cuticle development, proteolysis, and RNA transcription and translation in second stage larvae; peptidase inhibitor activity and nutrient reservoir activity in third stage larvae; cell and tissue morphogenesis and cell and tissue development in pupae; signal perception (a lot of genes implicated in light perception) and behaviors such as feeding, mating, and locomotion in adult flies. The analysis of gene expression related to parasitism revealed that 88 out of 480 peptidase genes, 110 out of 215 cuticle protein genes and 21 heat shock protein (hsp) genes showed a significant upregulation in the parasitic larval stages. Among peptidases, serine peptidases are actively involved in tissue degradation and nutrient acquisition during the parasitic processes of myiasis-causing larvae. Our investigation showed that up to 22.93% (47/205) of serine peptidase genes were highly expressed in the parasitic larval stages, but only 4.39% (9/205) in the pupal stage and 8.78% (18/205) in the adult stage. Interestingly, the expression of 2 antimicrobial peptide (AMP) genes, including 1 defensin and 1 diptericin upregulated in the parasitic second-stage larvae and 10 genes in the GO term "nutrient reservoir activity" is dominantly high in the parasitic third-stage larvae.

In the third article, we successfully isolated and characterized two pivotal sex-determination genes in *W. magnifica*, namely *Wmtra* (*W. magnifica transformer*) and *Wmtra2* (*W. magnifica transformer2*), whose orthologs have been utilized to develop genetic control approaches in several insect pests. The splicing of *Wmtra* transcripts exhibited a sex-specific pattern, with the

female variant encoding a fully functional protein and the male counterpart yielding a truncated, non-functional polypeptide, which is similar to blow flies such as *Lucilia cuprina* and *Cochliomyia hominivorax*. Furthermore, we identified the elements essential for the construction of a Cas9-based homing gene drive in *W. magnifica*. From sex-biased gene sets, a range of useful resources for the introduction site of the Cas9-gRNA cassette were provided. In addition, an examination of the set of genes exhibiting a bias toward females led to the identification of five potential candidate genes: *vasa* (*vas*), *nanos* (*nanos*), *bicoid* (*bcd*), *Bicaudal C* (*BicC*), and *innexin5* (*inx5*). The promoters of these genes can be used to drive Cas9 germline expression. Our further analysis suggested that the promoter from *nanos* is better suited than others due to its simple structure and successful use for this purpose in other Diptera. We have also identified six U6 RNA genes in *W. magnifica*; their promoters can serve as potential candidates for regulating gRNA transcription in Cas9-based homing gene drive.

Conclusions

In the first article, we successfully sequenced the genome of *W. magnifica* from a single fly specimen using the strategy of low DNA input library preparation. This approach holds promise for its applicability to small Diptera flies that are difficult to rear in a laboratory or are amenable to laboratory rearing but present difficulties in inbreeding. The genome of *W. magnifica* serves as a valuable resource for fundamental research in areas such as molecular biology, biochemistry, and genetics of *W. magnifica*. Moreover, it opens doors to numerous applications in comparative genomics, transcriptomics, functional genomics, population genetics, and notably, the development of novel control strategies. The second article offers an in-depth investigation into the dynamics of gene expression across the life cycle of *W. magnifica*, illuminating how the parasitic larvae upregulate specific gene expression to effectively parasitize their hosts. This study not only advances our understanding of the molecular-level parasitic life cycle of *W. magnifica* but also paves the way for innovative approaches to disrupt larval parasitism in the host. In the third article, we isolated and characterized essential elements for the construction of genetically modified strains in *W. magnifica*. This represents an important step towards the development of genetic control programs in *W. magnifica*. In summary, the thesis has shed light on the biology of *W. magnifica*, contributing to the development of novel control measures against the traumatic myiasis-causing flesh fly, *W. magnifica*.

5. GENERAL INTRODUCTION

5.1 What is myiasis?

Myiasis, a term derived from the Greek word for "fly", was initially coined by Hope in 1840 (Hope, 1840). It has since been defined as the infestation of live vertebrates (humans and/or animals) with dipterous larvae (maggots). These larvae, at least for a certain period, feed on the host's dead or living tissue, body substances, or ingested food (Zumpt, 1965).

Myiasis represents a prevalent parasitic infestation worldwide, with particular prevalence in warm tropical and subtropical regions (Hall et al., 2016; Noutsis and Millikan, 1994). Notably, a study revealed that myiasis ranks as the fourth most frequently encountered skin disease related to human travel in non-endemic areas (Caumes et al., 1995). In regions with a high prevalence of myiasis, these flies bring about significant economic losses and raise concerns about animal welfare in the field of animal husbandry (Francesconi and Lupi, 2012; Hall et al., 2016; Zumpt, 1965).

5.1.1 The classification of myiasis

Myiasis can be classified using two primary approaches: ecological and anatomical classifications (Francesconi and Lupi, 2012). The ecological classifications depend on the degree of a fly species' dependence on hosts and typically result in three categories: accidental, facultative, and obligatory myiasis (Francesconi and Lupi, 2012; Noutsis and Millikan, 1994; Scholl et al., 2019). The anatomical classification, on the other hand, focuses on defining the location within the host's body that is infested (Francesconi and Lupi, 2012; Zumpt, 1965) and generally divides myiasis into two main categories: cavitary myiasis and cutaneous myiasis (Francesconi and Lupi, 2012; Gour et al., 2018).

Accidental myiasis is also known as pseudomyiasis. The dipterous larvae typically do not have a parasitic nature. However, under specific and rare conditions, they may exhibit parasitic tendencies. For instance, this can occur when an animal ingests food that is contaminated with dipterous eggs or larvae, or when dipterous eggs or larvae invade through the host's anus (Noutsis and Millikan, 1994; Scholl et al., 2019). In facultative myiasis, dipterous larvae have

the capability to infest a host's tissue, but it is not their sole mode of development. The dipterous larvae have the flexibility to choose between parasitic or non-parasitic development, depending on the circumstances. (Noutsis and Millikan, 1994; Scholl et al., 2019). Obligatory myiasis refers to dipteran species whose larvae depend on infesting host tissue to complete their life cycle. In other words, the dipterous larvae involved are always parasitic; they require a living host for their development (Noutsis and Millikan, 1994; Scholl et al., 2019). Due to the dipterous larvae always parasitizing their hosts, obligatory myiasis is the most damaging.

Cavitary myiasis refers to dipterous larval infestations in the natural body cavities of hosts, including ophthalmic, ENT (Ear-nose-throat), intestinal and urogenital myiasis (Francesconi and Lupi, 2012; Gour et al., 2018).

Cutaneous myiasis is the infestation of the host's skin by dipterous larvae and is the most common type of myiasis (McGraw and Turiansky, 2008; Solomon et al., 2016). It can be further categorized into furuncular, migratory, and traumatic myiasis (McGraw and Turiansky, 2008; Robbins and Khachemoune, 2010; Solomon et al., 2016). Furuncular myiasis results in the formation of an erythematous furuncle-like nodule on the host's skin, which contains one or more maggots. Migratory myiasis, also known as creeping myiasis, occurs when dipterous larvae infest the host's skin and create serpentine or winding tunnels as they move beneath the skin's surface. Traumatic myiasis refers to dipterous larvae infesting open wounds in both humans and animals. Clinical manifestations of traumatic myiasis are characterized by visible dipterous larvae in and around the wound, pain, fever, chills, bleeding, and fistula formation (Francesconi and Lupi, 2012). Moreover, the presence of traumatic myiasis can increase the risk of secondary bacterial infection (Erdmann, 1987) of the wounds likely due to factors such as damaged skin, reduced immune responses, excretions, and exposure of wounds to a contaminated external environment. Traumatic myiasis is one of the most prevalent and severe forms of myiasis, resulting in significant economic losses and raising concerns about animal welfare in the field of animal husbandry (Hall et al., 2016).

5.1.2 The general life cycle of the myiasis-causing flies

According to the location of dipterous eggs hatching inside and outside adult female flies, the myiasis-causing flies are divided into two groups: oviparous and larviparous (Osorio-Pinzon et al., 2021). In the oviparous cases, the life cycle includes egg, larval, pupal, and adult stages

(Scholl et al., 2019). The life cycle begins with the egg stage. Adult female flies lay their eggs in suitable locations, typically either on or near hosts, where they await hatching. Once the eggs hatch, the larval stage begins. In this stage, the larvae (maggots) focus on feeding and growing. When the maggots reach a sufficient size and level of development, they enter the pupal stage. This stage is a resting stage, often located in the surrounding environment. The pupae ultimately emerge as adult flies. These adult flies are capable of flying, reproducing, and egg-laying in a new cycle of myiasis infestation.

Some myiasis-causing flies are larviparous, as they give birth to larvae rather than laying eggs (Scholl et al., 2019). This reproductive strategy ensures that the dipterous larvae can start their parasitic life more quickly.

In terms of obligatory myiasis, whether oviparous or larviparous species, the larvae employ diverse adaptive strategies to ensure their successful parasitism in a host. As the larvae enter the infestation site of hosts, they exhibit specific physiological activities and behaviors to adapt to the host's physiological environment, such as temperature, humidity, and acidity levels. Concurrently, to evade the host's immune response, larvae secrete substances that suppress the host's immune system. For instance, *Lucilia cuprina* secretes the blowfly larval immunosuppressive protein (BLIP) to suppress the host immune response (Elkington et al., 2009). In the case of Hypoderma spp., 2 peptidases, *hypodermin A* (*HA*) and *hypodermin B* (*HB*) are involved in coping with specific and non-specific host immune systems (Boulard, 1989; Chabaudie and Boulard, 1992; Moiré, 1998; Moiré et al., 1997; Pruett Jr, 1993). With the establishment of a suitable environment, larvae secrete more enzymes to break down host tissues for ingestion and nutrition acquisition for their development. Several studies on flies causing myiasis including *L. cuprina* and *Oestrus ovis* have shown that proteases, especially serine proteases, are the main enzymes that perform this function (Bowles et al., 1988; Casu et al., 1994; Sandeman et al., 1990; Tabouret et al., 2003).

5.2 The control of the myiasis-causing flies

Over the years, extensive research and practices have been carried out to address the issue of controlling myiasis-causing flies, including the direct targeting of either the larval or adult stages using chemical or biological agents, the manipulation of host immune responses, husbandry practices aimed at reducing the cues that attract myiasis-causing flies, breeding

programs to develop hosts with reduced susceptibility to myiasis, improved methods for predicting and detecting myiasis-causing flies, and the exploration of genetic manipulation techniques for eradicating or suppressing fly populations (Kotze and James, 2022; Sandeman et al., 2014). Among them, chemical and genetic methods are conspicuous in the control of myiasis-causing flies.

5.2.1. Insecticides

Insecticides have been applied to control myiasis for many years and, up until now, they remain the most effective and widely used approach. However, a major issue is that a large amount of insecticide usage may trigger myiasis-causing flies to develop resistance to insecticides. For example, blow flies, *L. cuprina* and *Cochliomyia hominivorax*, have exhibited resistance to a range of insecticides (Carvalho et al., 2009; da Silva and de Azeredo-Espin, 2009; Kotze and James, 2022; Tandonnet et al., 2022). In addition to resistance, the widespread use of insecticides has other adverse consequences such as harm to non-target organisms, residues in animal products with potential health risks to consumers, and safety concerns for those handling these chemicals (Ansari et al., 2014).

5.2.2 Genetic control

Genetic control, which targets the reproductive capacity of insect pest species, offers a promising approach to reducing population sizes to non-critical levels through the application of genetic manipulation techniques. This approach comes with several advantages, including species-specificity, reduced chemical usage, long-term effectiveness, precise targeting, and environmental friendliness.

5.2.2.1 Sterile Insect Technique (SIT)

Sterile Insect Technique (SIT) is one of the most well-known and successful genetic control strategies. Developed by Edward Knipling, Raymond Bushland, and their colleagues in the 1950s (Bushland et al., 1955; Knipling, 1955, 1959), SIT initially focused on combating the New World screwworm, *C. hominivorax*. Remarkably, this approach led to the complete eradication of screwworms from all of North and Central America over a 50-year period (Sandeman et al., 2014). In general, SIT includes large numbers of the target insect species

rearing in controlled environments such as insectaries or specialized facilities, sterilization of the insects using ionizing radiation that renders the insect incapable of reproducing but does not significantly affect their ability to mate or their overall fitness, and the release of sterile insects into the target area at a carefully calculated rate and pattern. After released sterile insects mate with their wild conspecifics in the target area, which results in sterile offspring, the overall population size of the target pest reduces to the desired levels through repeated mass releases of sterilized insects.

Limitations also exist in SIT. For example, the dose optimization of irradiation sterilization is necessary. Insects that receive too low doses are not sufficiently sterile and those that receive too high doses may not be competitive (Parker and Mehta, 2007). Establishing a balance between complete sterilization and maintaining the vitality and ability to mate for each insect species is challenging. Furthermore, some conventional SIT practices involve releasing both sexes but the concurrent release of females can result in competition between the released sterile females and wild females for mating opportunities with the sterile males (Knipling, 1959). This competition can necessitate a larger number of insects for population suppression. Field trials conducted on the Mediterranean fruit fly in Guatemala, have demonstrated that the male-only SIT release may be three to five times more effective than releasing both males and females in reducing the targeted populations (Rendón et al., 2004).

5.2.2.2 Conditional female lethal transgenic system

To overcome the limitations of SIT, the conditional female lethal transgenic sexing strains of myiasis-causing flies have been developed in *L. cuprina* (Li et al., 2014) and *Cochliomyia hominivorax* (Concha et al., 2016). The respective strain contains a single-component tetracycline-repressible system, which includes a marker gene and a sex-specific *tetracycline-dependent transactivator* (*tTA*) overexpression cassette made up of a tet operator (tetO), a promoter, an intron from the *tra* gene, and a *tTA* gene. The *tTA*, a fusion of the DNA binding domain from the E. coli tet repressor and the viral VP16 transcription activation domain (Gossen and Bujard, 1992), is a key component. The overexpression of *tTA* leads to lethality, possibly due to phenomena like "transcriptional squelching" or interference with ubiquitin-dependent proteolysis (Gong et al., 2005). The promoter is responsible for driving the expression of the *tTA*. The *tra* gene plays a crucial role in sex determination and exhibits sex-specific splicing in *L. cuprina* (Concha and Scott, 2009) and *C. hominivorax* (Li et al., 2013),

resulting in female flies encoding a full-length TRA protein, while male flies produce a truncated, nonfunctional polypeptide due to in-frame stop codons. When the sex-specific *tra* intron is introduced into the *tTA* gene, it enables the sex-specific transcription of the *tTA* within the system. In the presence of tetracycline in the diet, tetracycline can bind to the produced *tTA*, allowing both sexes to survive and rear in a controlled facility. Conversely, in the absence of tetracycline, the sex-specific *tra* intron ensures that only female flies can encode functional tTA protein. This innovative approach effectively achieves the aim of allowing only male flies to survive.

In the case of the single-component system described above, female lethality typically occurs at the third larval or pupal stages (Concha et al., 2016; Li et al., 2014), and further cost savings could be achieved if females die at the embryonic or first larval stages. To address this, two-component female embryonic lethal systems have been developed in *L. cuprina* (Yan and Scott, 2020) and *C. hominivorax* (Concha et al., 2020). The two-component system comprises a driver construct and an effector construct. The driver construct includes an early-acting embryonic promoter that controls the expression of the *tTA* gene. The effector construct contains a pro-apoptotic gene that is regulated by the *tTA*. To make this system female-specific, the first sex-specific *tra* intron is inserted after the start codon of the pro-apoptotic gene. As in the single-component system, only females die when tetracycline is absent from the diet, while both females and males remain fully viable if tetracycline is introduced into the insect diet (Figure 1). In the two-component system, since the *tTA* expression is regulated by promoters active during the embryonic or first larval stages, females can be selectively eliminated at the early developmental stages, which can save the costs of producing female flies. Furthermore, after released males mate with wild females in the field, the female offspring will die before they can cause any harm to animals.

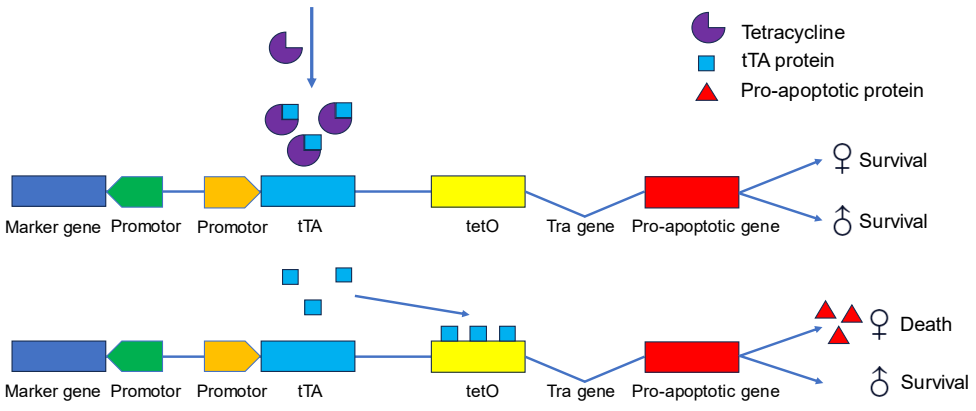

Figure 1. The schematic diagram of the genetic constructs of a two-component transgenic female early lethal system. The two-component system consists of a Driver construct that expresses *tTA* under the control of a promoter, and an Effector construct composed of tetO, the first intron of the *tra* gene, and the pro-apoptotic gene. In the upper part, due to the presence of tetracycline in the diet, tetracycline can bind to *tTA*, leading to a strong inhibition of the binding of *tTA* to tetO. This inhibition prevents the expression of the pro-apoptotic gene. Therefore, both female and male flies can survive. In the lower part, when tetracycline is lacking in the diet, *tTA* can bind to tetO. Due to the presence of the sex-specific *tra* intron, only female flies can express the pro-apoptotic gene. Consequently, female flies die, while male flies do not express the pro-apoptotic gene and therefore survive. *tTA*: *Tetracycline repressible transactivator*; tetO: Tet operator; *tra*: The *transformer* gene.

5.2.2.3 Gene drives

Gene drives are another powerful genetic engineering technique aimed at reducing populations of harmful insect pests (Bier, 2022). A gene drive system dependent on CRISPR-Cas9 technology has been developed in several mosquito species (Carballar-Lejarazú et al., 2020; Gantz et al., 2015; Hammond et al., 2016; Kyrou et al., 2018). Recently, the gene drive has also been developed for the suppression of *Drosophila suzukii* (Yadav et al., 2023), an invasive agricultural insect pest of soft-skinned fruits. In concept, the Cas9-based homing gene drive is a mechanism that biases the inheritance of a particular gene or genetic trait within a population using CRISPR-Cas9 technology compared to the probability of one half expected under Mendelian inheritance (Bier, 2022). At the molecular level, a cassette containing an RNA-guided Cas9 endonuclease and its guide RNA (gRNA) is inserted into a target gene. Among them, the Cas9 endonuclease is controlled by a germline-specific promoter, and gRNA is promoted by a U6 promoter. After the insects bearing a gene drive construct mate with the wild insects, the Cas9 endonuclease generated by the construct is guided by the gRNA to cleave a target site on the chromosome in wild insects. With the chromosome containing the Cas9 endonuclease and gRNA sequences as a repair template, the double-strand break is repaired by homology-directed repair (HDR) as part of the DNA repair process, leading to the "homing" of gene drive (the conversion of the drive-bearing locus from heterozygosity to homozygosity) and transmitting to offspring (Figure 2). Since the targeted gene is essential for the viability or fertility of female insects, female offspring either do not survive or lose the ability to reproduce,

thus achieving the suppression of insect pests after several generations.

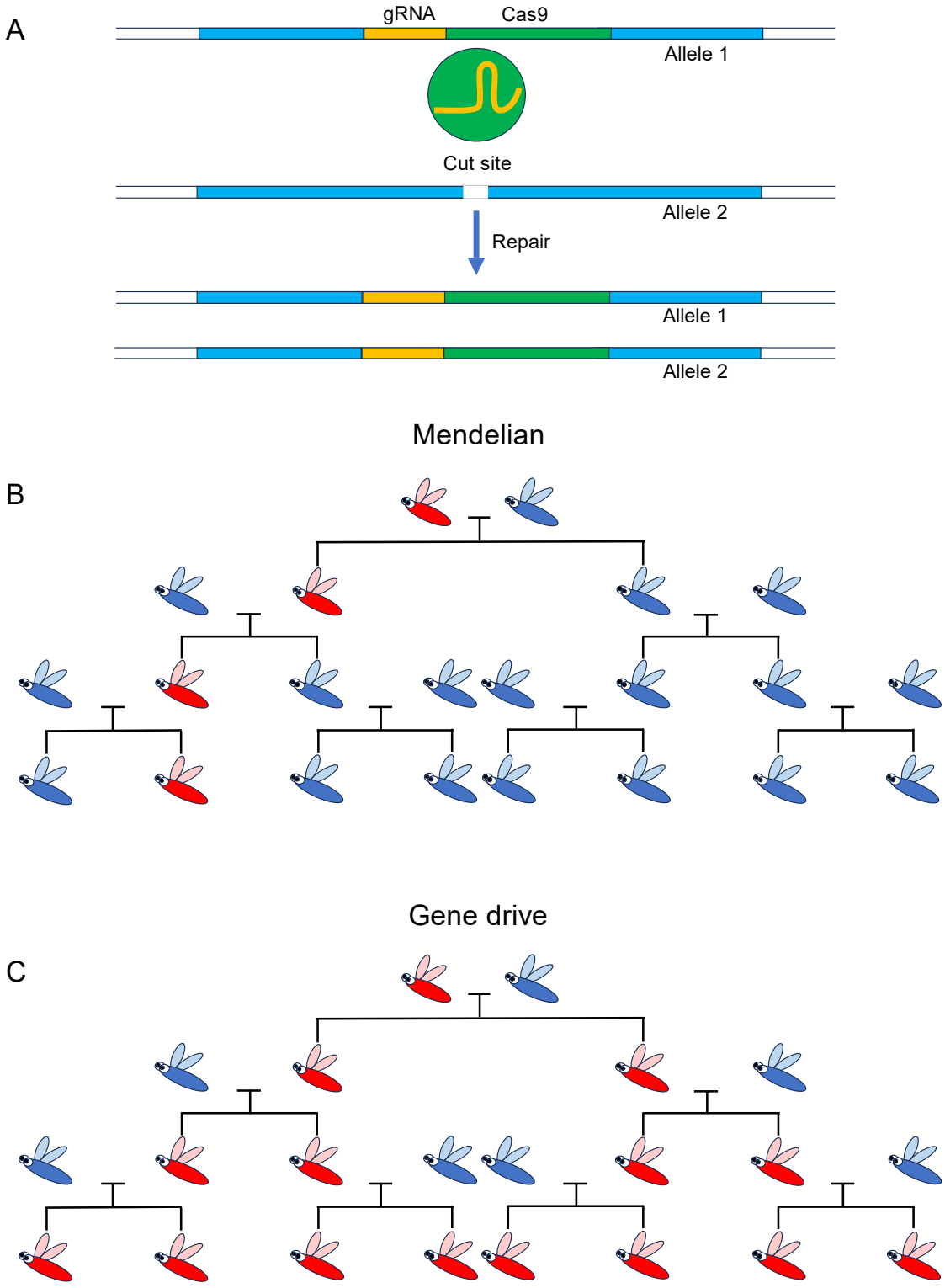

Figure 2. The schematic diagram of the Cas9-based homing gene drives. **A**. A cassette encoding gRNA (yellow) and Cas9 (green) is inserted into the genome, and the Cas9 + gRNA

cassette then directs the cleavage of homologous chromosomes in the germline and is copied into the DNA breaks by homology-dependent repair (HDR), so that nearly all progeny (~99%) inherit the "gene drive" cassette. **B** and **C**. The comparison of Mendelian (B) and gene-driven (C) inheritance patterns. In each case, a small number of transgenic individuals (red) are introduced into a large wild-type (WT) population (blue).

5.3 Wohlfahrtia magnifica

W. magnifica (Schiner, 1862; Diptera: Sarcophagidae), also known as the spotted flesh fly, is one of the most important obligatory traumatic myiasis-causing dipteran flies (Francesconi and Lupi, 2012; Scholl et al., 2019). The species can infest a wide range of mammals, including camels, sheep, goats, horses, deer, cattle, and even humans (Diakakis et al., 2006; Farkas et al., 1997, 2009; Gaglio et al., 2011; Hall et al., 2009b; Giangaspero et al., 2011; Ruiz Martinez and Leclercq, 1994; Sotiraki et al., 2010; Valentin et al., 1997; Yasuda, 1940). Usually, the genitalia of both males and females are the primary infestation sites (Figure 3), accounting for approximately 87% of cases (Farkas et al., 1997; Ruiz Martinez et al., 1987, 1991). The impact on infested hosts includes reproductive disorders, lameness, blindness, and potentially fatal consequences if left untreated, raising concerns about animal welfare, health issues, and substantial economic losses in livestock production (Farkas et al., 1997; Hall & Wall, 1995).

Figure 3. Parasitic larvae of *W. magnifica* in the vagina of a Chinese Bactrian camel included in the samples collection of this thesis.

5.3.1 The life cycle of W. magnifica

W. magnifica exhibits its highest activity levels during the warmer summer months, typically from May to September each year (Cruz et al., 1996). *W. magnifica* is a larviparous species, therefore, the life cycle of *W. magnifica* comprises three stages, including two free stages (the pupal and adult stages) and an obligatory parasitic larval stage (Cruz et al., 1996). In the active season, female *W. magnifica* are attracted to wounds or natural body orifices of their hosts and deposit the first-stage larvae directly around or in the infestation sites in the hosts. These newly deposited larvae feed on the host's cutaneous and underlying tissues for 4-8 days, progressing from the first larval stage to the third larval stage. If not treated promptly, the feeding activity can lead to significant tissue destruction. As this feeding behavior continues, it attracts more females to arrive and deposit more larvae, causing the wounds to further enlarge and deepen. The fully developed third-stage larvae exit the wounds, fall to the ground, burrow into the soil and, pupate. Adult flies emerge from the pupae within 11-15 days. About 4 days after mating, female flies seek a new host to deposit their first-stage larvae and start a new life cycle (Cruz et al., 1996). Additionally, during winter, the pupae enter a diapause state in response to the cooler weather.

5.3.2 The epidemiology of W. magnifica

W. magnifica is primarily distributed in Europe, Asia, and Africa (Diakakis et al., 2006; Farkas et al., 1997, 2009; Gaglio et al., 2011; Hall et al., 2009b; Giangaspero et al., 2011; Ruiz Martinez and Leclercq, 1994; Sotiraki et al., 2010; Valentin et al., 1997; Yasuda, 1940). In Europe, *W. maginifica* is a significant myiasis-causing insect pest in sheep. In Romania, a high infestation rate of 80% to 95% was observed in sheep, resulting in a 20% mortality rate among newborn lambs (Lehrer et al., 1988). In Spain, Remesar et al. (2022) conducted a study on traumatic myiasis in 73,683 sheep from 122 flocks in Albacete Province. The results indicated an overall flock prevalence of 95.9%, with individual prevalence at 7.1%, and *W. magnifica* was the only species found in all positive cases, except for one case which was infested with *Lucilia sericata*. In Hungary, a survey conducted between 1992 and 1995 found that 17.6% (774/4388) of the examined sheep were infested with active traumatic myiasis, and except for five cases caused by *L. sericata*, *W. magnifica* was the only species identified in these wounds (Farkas et al., 1997). In China and Mongolia, the infested animals include sheep and Bactrian camels. Li et al. (2019b) investigated six sheep flocks (comprising 2261 sheep) in Gansu, China,

and except for one flock where no cases of infestation were detected, the prevalence of the remaining five flocks ranged from 4.5% to 5.00%. A survey of 2038 Bactrian camels in the Inner Mongolia Autonomous Region of China showed an overall manifestation of myiasis of 26.6% (Liu et al., 2022). In Mongolia, Valentin et al. (1997) examined 1676 Bactrian camels from 45 selected herds in six different areas of the Chatanbulag Sumon in the Eastern Gobi district and the result showed *W. magnifica* infestation rates between 8% and 10%. It is important to note that the *W. magnifica's* impact on wildlife can be particularly severe, as this parasite is difficult to control within wildlife populations, leading to devastating consequences. For example, Yan et al. (2019) reported traumatic myiasis infestation in endangered Przewalski's horses caused by *W. magnifica*. Another example is the critically endangered wild camels (*Camelus ferus*) living in Inner Mongolia and China wild camel (approximately 600 surviving in China and 450 in Mongolia), where *W. magnifica* infestation is prevalent.

5.3.3 The progress in the study of W. magnifica

So far, research on *W. magnifica* has been quite limited. The primary focus of existing studies revolves around epidemiological investigations (Farkas et al., 1997; Liu et al., 2022; Remesar et al., 2022) and case reports of *W. magnifica* infestations (Beyhan et al., 2017; Çiftçioğlu et al., 1996; Maurya et al., 2012). Several research endeavors have also explored laboratory rearing (Cruz et al., 1996, 1998; Farkas et al., 2005) and morphological observation (An et al., 2019; Li et al., 2020) of *W. magnifica*, as well as bacteria isolation from *W. magnifica* (Toth et al., 2006, 2008; Tóth et al., 2001). At the molecular research level, the limited work primarily concentrated on mitochondria. For example, Zhang et al. (2016) sequenced the mitogenome of *W. magnifica*. Mitochondrial genes, such as the cytochrome b gene, have been utilized in phylogenetic analysis related to *W. magnifica* (Hall et al., 2009a; Marangi et al., 2016). As for the control of *W. magnifica*, limited research involved the investigation of insecticides, including ivermectin and moxidectin (Farkas, et al., 1996), dicyclanil (Sotiraki et al., 2005), as well as cypermethrin and doramectin (Sotiraki et al., 2003). Furthermore, a study explored the potential of naturally derived, plant-based formulations for managing *W. magnifica's* infestations, showing promising results (Carnevali et al., 2019).

5.4 The sequencing technology applications in insects

5.4.1 Genome

Genomes are the starting point for a number of fundamental and applied studies. In insects, genomic research contributes to the understanding of basic biology (e.g., the mechanisms of gene regulation in development, growth, reproduction, etc.), evolutionary history and relatedness, and insect pest management (e.g., the identification of the genes associated with insect resistance and targets for novel insecticides and vaccines) as well as the development of genetically modified strains and gene drive systems. Previously, high-quality genome assemblies of insects were obtained from reads generated using short-read sequencing technology, with Illumina being the widely adopted platform due to its cost-effectiveness and high accuracy. However, its limitation in generating short reads (<250 bp) often results in genome assemblies with a low contig N50 length. Long-read sequencing technologies are capable of generating relatively long read fragments that can address complex structures and highly repetitive regions in a genome. This is important for the correct and high-quality genome assembly. PacBio Single Molecule Real-Time (SMRT) and Oxford Nanopore are currently the most widely used long-read sequencing technologies. Recently, the genomes of several insects have been publicly reported using PacBio/Oxford or a combination of Illumina and PacBio/Oxford (Anstead et al., 2015; Li et al., 2019a; Scott et al., 2020). However, long-read sequencing technologies have high error rates. In addition, whether it is a short-read or long-read strategy, due to the high degree of heterozygosity in most insects and their small size, it is necessary to pool together many inbred individuals to enhance genetic homogeneity and to obtain sufficient DNA for these strategies. However, for most insects, it is quite difficult or impossible to produce inbred strains. Therefore, methods that bypass the need to pool organisms are crucial for facilitating the creation of reference-quality genomes, especially for insects captured in the wild.

Kingan et al. (2019) pioneered a modification to the PacBio SMRTbell library construction protocol without DNA shearing and size selection. They achieved a high-quality de novo genome assembly using a single *Anopheles coluzzii* mosquito, significantly reducing DNA input requirements and eliminating the need for organism pooling. Recently, PacBio has evolved to HiFi long reads. These HiFi reads provide a base-level accuracy of 99.9%, on par with Illumina short reads and Sanger sequencing. Combining the modified low DNA input library preparation with HiFi long-read sequencing technology, highly accurate genome assemblies can be obtained for a number of insects that are small in size and also have difficulty in obtaining inbred strains.

5.4.2 Transcriptome

Transcriptomes serve as powerful tools for investigating the dynamics of gene expression levels and structural variations. In insect research, transcriptomics is involved in a range of analyses such as the quantification of the gene expression specific to tissues and developmental stages, response to external pressures under different environmental conditions, interactions between insects and host plants or host animals, and interactions between insects and pathogens and the mechanisms of disease transmission in disease-carrying insects, as well as the identification of structural variations. Until now, transcriptomes of numerous insects have been sequenced, including blow flies (Anstead et al., 2015; Scott et al., 2020). Currently, the Illumina sequencing platform is the most commonly used and widely accepted in the investigation of gene expression levels because of its high throughput, high accuracy, and relatively low cost. In addition, it can capture and measure almost all RNA types in a sample, including mRNAs, rRNAs, tRNAs, and non-coding RNAs, and also can detect low-abundance RNA molecules. However, Illumina-based RNA-seq sequencing platforms typically yield short read lengths, presenting a challenge for detecting complex gene structures, long transcripts, or alternatively spliced transcripts. Addressing these challenges, the Iso-Seq method sequences (Gonzalez-Garay, 2016) entire cDNA molecules, reaching lengths of up to 10 kb or more, without the need for the bioinformatics transcript assembly. Therefore, this technology excels in characterizing alternative splicing (AS) events (e.g., alternative start sites, end sites, intron retentions, and exon-skipping events), finding gene fusions, identifying allele-specific isoforms, detecting differentially expressed isoforms and isoform switching events, and predicting the functional impact of novel isoforms through open reading frame (ORF) prediction.

5.5 The aims of the PhD thesis

The aims of my PhD thesis were to address research gaps related to *W. magnifica*. These aims include:

i) Sequencing, annotating, and characterizing the genome of *W. magnifica* and exploring the evolutionary relationships between *W. magnifica* and other dipteran species. The availability of the genome of *W. magnifica* will facilitate various fundamental and applied research related

to *W. magnifica*.

ii) Investigating the complex gene expression dynamics throughout the life cycle of *W. magnifica* to understand how the parasitic larvae regulate specific gene expressions to effectively parasitize their hosts. This research can serve as a guide for identifying potential targets for vaccines or insecticides designed to disrupt the larvae's ability to establish themselves on or in hosts.

iii) Isolating and characterizing the essential components for the development of genetically modified strains, including the transgenic sexing strains and the CRISPR/Cas9-based homing gene drive systems for *W. magnifica*. This marks a significant step towards establishing genetic control programs to combat the infestation of *W. magnifica*.

6. PUBLICATIONS

6.1 Article 1

Jia, Z., Hasi, S., Vogl, C., & Burger, P. A. (2022). Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasi-causing fly of warm-blooded vertebrates. *Molecular Ecology Resources*, 22(7), 2744-2757. https://doi.org/10.1111/1755-0998.13654
Impact factor: 7.7

The data created within the article has been deposited at the following locations:

The genome assembly of *W. magnifica* was deposited as a BioProject under the accession number PRJNA778059. The PacBio HiFi sequence reads have been deposited at NCBI under the accession number SRR16848117. The transcriptome data has been deposited in the Sequence Read Archive (SRA), including 3 first-stage larvae (SRR18178228, SRR18178229, SRR18178230), 3 second-stage larvae (SRR18178225, SRR18178226, SRR18178227), and 3 adult flies (SRR18178222, SRR18178223, SRR18178224). In addition, the assembly and annotation of the *W. magnifica* genome are also available on Dryad (https://doi.org/10.5061/dryad.qfttdz0j8). Supplementary materials, including Table S1 to Table S6 and Data S1, can be found immediately following the main text of this article in this thesis.

Received: 7 November 2021 | Revised: 1 May 2022 | Accepted: 17 May 2022

DOI: 10.1111/1755-0998.13654

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates

Zhipeng Jia[1] | Surong Hasi[2] | Claus Vogl[3] | Pamela A. Burger[1]

[1]Department of Interdisciplinary Life Sciences, Research Institute of Wildlife Ecology, University of Veterinary Medicine Vienna, Vienna, Austria

[2]Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Hohhot, China

[3]Department of Biomedical Sciences, Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna, Vienna, Austria

Correspondence
Pamela A. Burger, Department of Interdisciplinary Life Sciences, Research Institute of Wildlife Ecology, University of Veterinary Medicine Vienna, Savoyenstrasse 1, Vienna 1160, Austria.
Email: pamela.burger@vetmeduni.ac.at

Funding information
Austrian Science Fund, Grant/Award Number: P29623-B25; Chinese Scholarship Council, Grant/Award Number: 201909150004

Handling Editor: Kin-Ming (Clement) Tsui

Abstract

Wohlfahrtia magnifica is a pest fly species, invading livestock in many European, African and Asian countries, and causing heavy agroeconomic losses. In the life cycle of this obligatory parasite, adult flies infect the host by depositing the first-stage larvae into body cavities or open wounds. The feeding larvae cause severe (skin) tissue damage and potentially fatal infections if untreated. Despite serious health detriments and agroeconomic concerns, genomic resources for understanding the biology of *W. magnifica* have so far been lacking. Here, we present a complete genome assembly from a single adult female *W. magnifica* using a Low-DNA Input workflow for PacBio HiFi library preparation. The de novo assembled genome is 753.99 Mb in length, with a scaffold N50 of 5.00 Mb, consisting of 16,718 predicted protein-encoding genes. Comparative genomic analysis revealed that *W. magnifica* has the closest phylogenetic relationship to *Sarcophaga bullata* followed by *Lucilia cuprina*. Evolutionary analysis of gene families showed expansions of 173 gene families in *W. magnifica* that were enriched for gene ontology (GO) categories related to immunity, insecticide-resistance mechanisms, heat stress response and cuticle development. In addition, 45 positively selected genes displaying various functions were identified. This new genomic resource contributes to the evolutionary and comparative analysis of dipterous flies and an in-depth understanding of many aspects of *W. magnifica* biology. Furthermore, it will facilitate the development of novel tools for controlling *W. magnifica* infection in livestock.

KEYWORDS
comparative genomics, de novo assembly, fly species, gene family, low DNA input, PacBio HiFi sequencing, positively selected genes

1 | INTRODUCTION

The obligate parasitic spotted flesh fly, *Wohlfahrtia magnifica* (Diptera: Sarcophagidae), is globally distributed, ranging from North Africa, through eastern and southwestern Europe extending to northeast Asia (Farkas et al., 1997; Gaglio et al., 2011; Giangaspero et al., 2011; Hall et al., 2009; Ruiz Martinez & Leclercq, 1994; Sotiraki et al., 2010; Valentin et al., 1997; Yasuda, 1940). It is a major myiasis-causing fly and can infect live mammals (Schnur et al., 2009), especially livestock, such as horses (Farkas & Képes, 2001; Yan et al., 2019), sheep

(Dehghani et al., 2014; Farkas et al., 1996), camels (Moshaverinia et al., 2013; Valentin et al., 1997), and even occasional reports of infected humans (ÇiftÇio et al., 1996; Kokcam & Saki, 2005). Similar to other dipterous larvae in the family Calliphoridae that infect livestock, e. g., *Lucilia cuprina* (Anstead et al., 2015) and *Cochliomyia hominivorax* (Scott et al., 2020), adult *W. magnifica* females seek their hosts and lay their larvae on the predisposed skin of the genitalia or open wounds. The larvae feed on the host's tissues for development. As a result, the numerous bites of the larvae enlarge the wound and lead to severe tissue damage within only a few days. This may cause emaciation, reduction of productivity, reproductive disorders, and if untreated, heavy infections possibly leading to death (Farkas et al., 1997; İpek et al., 2012; Martinez et al., 1987).

Economic loss owing to *W. magnifica* infection in livestock worldwide is considerable (Hall, 1997; Ruiz-Martínez et al., 1987). In China, field surveys show that approximately 20% of female Bactrian camels are infected each year. In Europe, the individual prevalence of *W. magnifica* ranges from 0.7% to 95%, especially in Spain and Italy (Hall & Farkas, 2000; Sotiraki et al., 2012). For example, Remesar et al. (2022) surveyed 122 flocks of 73,683 sheep in Albacete Province of Spain, and the result showed that 90% of flocks were infected and the prevalence of the individuals was 7.1%. Over many years, different control methods have been employed, mainly applying insecticides to control larval development (Farkas et al., 1996; Giangaspero et al., 2011; Sotiraki et al., 2005). However, insecticides cannot prevent the infection/reinfection and may cause undesired side effects, for example, local corrosion and necrosis of skin and tissue at the site of infection, leading to festering wounds, prone to secondary infection (Hall & Farkas, 2000). In addition, the excessive use of insecticides could cause the emergence of insecticide resistance. Therefore, the existing approaches for the control of infection with the *W. magnifica* are still limited and not applicable for long-term use. New alternative control tools, such as vaccine-based approaches, sterile insect technique (SIT), development and discovery of new drugs, or other genetic, immunological or chemical control strategies are greatly needed.

Until today, most research on *W. magnifica* has focused on epidemiological studies (Ruiz Martínez et al., 1993), morphological observations (Li et al., 2020; Szpila et al., 2014; Yasuda, 1940), and investigation of life history (Cruz et al., 1996), while little is known about the host's immune response, parasite–host interaction, or regulatory mechanisms on the level of molecular biology. A high-quality genome is a fundamental resource for understanding many aspects of the developmental and reproductive biology, physiology and biochemistry as well as complex pathogenic mechanisms of *W. magnifica* or for developing novel control methods preventing fly invasion to livestock. However, due to the presence of insect-highlighting features such as high polymorphism, high-quality genome assembly is difficult to obtain (Richards & Murali, 2015). A representative obstacle is that repetitive sequences or polymorphic regions cannot be straddled well, leading to a fragmented genome assembly with lower contig N50 lengths. As long-read technologies have the inherent advantages of spanning polymorphic regions, repetitive sequences and transposable elements (TE) (Richards & Murali, 2015), at present, more and more insect genome projects are a combination of continuous long reads from PacBio/Nanopore and short reads from Illumina (Meng et al., 2020; Ren et al., 2021; Ye et al., 2021). Using the strategy, to date, a number of full and draft genomes of agricultural pests are completed and publicly available, such as Mediterranean fruit fly (*Ceratitis capitata*) (Papanicolaou et al., 2016), sheep blow fly (*L. cuprina*) (Anstead et al., 2015), and the New World screwworm fly (*C. hominivorax*) (Scott et al., 2020). This strategy requires a sufficiently high quantity of DNA for library preparation for insects with small physical sizes, which can normally be obtained from the time-consuming rearing of inbred lines. However, some attributes of *W. magnifica* have severely hindered a high-quality genome assembly so far as insufficient DNA quantities obtained from a single physically small adult fly have posed a major problem. One way to overcome this challenge would be to pool inbred individuals to obtain sufficient DNA for library preparation. While this method works well for some organisms that can be inbred, such as *Drosophila melanogaster* (Adams et al., 2000), unfortunately, *W. magnifica* is notoriously difficult to rear (in vitro rearing) (Farkas et al., 2005). In the laboratory, *W. magnifica* has a very high mortality rate, which renders the inbreeding strategy unfeasible. For example, researchers used dead animals or their tissues as larval diets to rear *W. magnifica*, unfortunately without success (Ruiz Martinez et al., 1992; Soler Cruz et al., 1998). Another attempt would be to feed on the artificial diet and the results showed 64%–98% mortality in the larval stage, 61%–100% mortality in the pupal stage, and only a maximum of 6% were successfully reared from the first stage larvae to the adult stage (Farkas et al., 2005).

The availability of the high fidelity (HiFi) library preparation workflow from low-DNA input (Kingan et al., 2019) has improved this situation. Compared to the standard HiFi library preparation of PacBio, which requires relatively large DNA amounts (a minimum input of 5 μg high-molecular-weight genomic DNA is recommended for the Sequel II systems), this workflow significantly reduces DNA requirement. As a result, very small amounts of genomic DNA (>400 ng for the Sequel II system) isolated from a single insect can produce sufficient amounts of sequencing data for a high-quality genome assembly of up to 1 Gb using only one SMRT cell combined with the circular consensus sequencing (CCS) mode of the PacBio Sequel II system. This also avoids time-consuming inbreeding and pooling requirements.

Here, we report a high-quality and accurate genome assembly of *W. magnifica* with the size of 753.99 Mb, including complete annotation. This resource can assist in enlightening the genetic mechanisms of *W. magnifica* and eventually in developing applications to control this invasive fly species. Finally, we performed comparative genome analyses with other dipterous flies, allowing us to gain a better understanding of the molecular evolution of *W. magnifica*.

2 | MATERIALS AND METHODS

2.1 | Sample collection

The study site is located at the Camel Culture Base in Siziwang Banner, Ulanqab City (Inner Mongolia Autonomous Region, China). The sample collection was performed in the frame of veterinary

health monitoring and treatment of Bactrian camels with *W. magnifica* infection. The first, second, and third stages of *W. magnifica* larvae were obtained from the genitalia of an infected female Bactrian camels in a lie-down position, of which the first and second stages were used for RNA extraction (Z. Jia, S. Hasi, C. Vogl, P. A. Burger, unpublished data). For adult fly rearing, the third-stage larvae were placed in preprepared foam boxes with the local soil. After the collected larvae had burrowed into the soil, the foam boxes were brought to the laboratory and placed in a dry place for hatching. Around 18–20 days, when all the third-stage larvae had emerged as adult flies, the adult females with the largest relative body size were selected and frozen at −80°C until DNA and RNA extraction.

2.2 | DNA isolation and sequencing

High-molecular-weight genomic DNA was extracted from a single adult female *W. magnifica*. The quantity of extracted DNA was measured using an Invitrogen Qubit 3.0 fluorometer (Thermo Fisher Scientific) and Nanodrop NC2000 (Thermo Fisher Scientific), and the integrity of extracted DNA was estimated on a 1.2% agarose gel to check for any degradation. With approximately 1.215 μg DNA isolated from a single female *W. magnifica* (Table S1), a 10 kb HiFi library was prepared following the procedure and recommendations of the kit: Preparing HiFi Libraries from low DNA input using SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences). In short, genomic DNA was sheared to average size distribution of 10 kb using g-TUBEs (Covaris) and subsequently purified. Purified DNA fragments were added to the enzyme reaction tubes and incubated at 37°C for 15 min to remove single-strand overhangs followed by the addition of repair mix and incubation at 37°C for 30 min to repair the damage within the DNA backbone. After DNA damage repair, the ends of the double-stranded fragments were polished and subsequently tailed with an A-overhang by adding End Prep Mix and incubating at 20°C for 10 min and then at 65°C for 30 min. Ligation with T-overhang SMRTbell adapters occurred at 20°C for 60 min, after which the AMPure PB beads (Pacific Biosciences) were employed to purify the SMRTbell library. Due to the presence of short fragments after the first purification step, the library was size-selected with AMPure PB beads (Pacific Biosciences) to remove SMRTbell fragments less than 3 kb. Subsequently, the size distribution and quantity of the SMRTbell library to be sequenced were measured using Invitrogen Qubit 3.0 Fluorometer (Thermo Fisher Scientific) and Agilent 2100 Bioanalyser (Agilent Technologies). The final SMRTbell library was sequenced on the Pacbio Sequel II system with a single SMRT Cell 8M.

2.3 | RNA extraction, library preparation, sequencing and data filtering

To assist the assessment of genome assembly and genome annotation, we sequenced transcriptome data from the first- and second-stage larvae and adult flies of *W. magnifica* with three replicates for each sample. Total RNA was extracted from each sample. RNA quality was examined by agarose gel electrophoresis and Agilent 2100 Bioanalyser (Agilent Technologies). Library preparation followed the instructions of TruSeq Stranded mRNA LT Sample Prep Kit (Illumina). Briefly, mRNA was enriched by binding to poly-A on mRNA with Beads containing oligo-dT, and the enriched mRNA was interrupted to a 200–300 bp fragment. Then, the fragmented RNA was used as a template for reverse transcription to the first-strand complementary DNA (cDNA) synthesis, followed by the synthesis of second-strand cDNA, which uses the first-strand cDNA as a template. After that, synthetic double-stranded cDNA was end-repaired, poly (A) added, and ligated to Illumina sequencing adapters. The ligation products were first purified by removing the free adaptor and the fragment without the attached adaptor, and next amplified by PCR using specific primers. Finally, the prepared libraries were sequenced on an Illumina NovaSeq platform. To obtain high-quality clean reads, the raw paired-end reads were trimmed by removing adapter sequences and low-quality reads using bbmap (https://sourceforge.net/projects/bbmap/).

2.4 | Genome assembly

To correct sequencing errors and generate highly accurate consensus reads, we converted raw reads into circular consensus sequences (CCS; hereafter HiFi sequences) using the program ccs version 5.0.0 with default settings (https://github.com/PacificBiosciences/ccs). Next, we used IceCreamFinder version 38.84 (https://sourceforge.net/projects/bbmap/) to filter out and/ or trim HiFi sequences with inverted repeats and remaining adapter sequences with default settings. Then, we filtered the resulting HiFi reads for potential bacterial contamination using SendSketch version 38.87 (https://sourceforge.net/projects/bbmap/) to send a reduced representation of the trimmed/ filtered HiFi reads against drafts from the NCBI nucleotide database inspecting up to 1000 records in the results. We used NCBI data sets version 10.9.0 to retrieve matching bacterial genome sequences and Seal version 38.87 (https://sourceforge.net/projects/bbmap/) with k = 31 and minkmerfraction = 0.5 to assign and remove HiFi sequences with at least 50% of each HiFi sequence's 31-mers matching the bacterial genomes. For the genome assembly based on the filtered HiFi sequences, the official PacBio software for HiFi genome assembly, the Improved Phased Assembler (IPA) version 1.3.2 (https://github.com/PacificBiosciences/pbbioconda/wiki/Improved-Phased-Assembler), was employed with default settings.

To assess the completeness of the genome assembly, we applied the genome mode of the Benchmarking Universal Single-Copy Orthologues (Busco, version 4.0.6) (Simão et al., 2015) and searched for conserved single-copy genes belonging to the core gene sets of diptera_odb10 (Creation date: 2020-08-05, number of species: 56, number of BUSCOs: 3285). In addition, we compared the BUSCO scores between the *W. magnifica* genome assembly and other dipterous flies, including *Lucilia cuprina* (ASM118794v1), *Musca domestica* (MdomA1), *Stomoxys calcitrans* (ScalU1), *Glossina morsitans*

(GmorY1), *Drosophila melanogaster* (BDGP6.32), *Mayetiola destructor* (Mdes_1.0), *Aedes aegypti* (AaegL5), *Anopheles gambiae* (AgamP4) from Ensembl Metazoa and *Sarcophaga bullata* (GCA_005959815.1) from NCBI.

2.5 | Annotation of repetitive sequences

We soft-masked (converted uppercase to lowercase bases) the genome assembly by generating a species-specific repeat library with RepeatModeler version 2.0.1(http://www.repeatmasker.org/RepeatModeler/) using -engine ncbi and -LTRStruct. The repeat library from RepeatModeler was filtered to remove known UniProt/SwissProt version 2020_05 proteins using ProtExcluder version 1.1 (Campbell et al., 2014). We then used RepeatMasker version 4.1.1 (http://www.repeatmasker.org/) with the options "-xsmall -a" and with the species-specific repeat library to identify repetitive sequences.

2.6 | Gene annotation

We annotated the genome assembly with Braker version 2.1.5 (Hoff et al., 2019), Augustus version 3.3.3 (Stanke et al., 2004), and GeneMarkES version 4.6.3 (Lomsadze et al., 2005). We used proteins from Arthropoda v100_odb10 (Kriventseva et al., 2019), RNA-Seq alignments made between RNA-Seq libraries aligned to the genome with hisat2 version 2.2.1 (Kim, Nam, et al., 2019; Kim, Paggi, et al., 2019) using --max-intronlen 100,000 and --dta. For BRAKER we used the softmasking, etpmode, and the following augustus settings: --alternatives-from-sampling = true –minexonintronprob = 0.2 –minmeanexonintronprob = 0.5 –sample = 100 –maxtracks = 3 –temperature = 2. Then, we employed Maker version 3.01.03 (Cantarel et al., 2008) to merge the annotations by Augustus and GeneMark using the hints-file.gff produced by Braker as the protein_gff passed to Maker and the concatenated augustus.hints.gtf and GeneMark-ETP's genemark.f.multi_anchored.gtf filtered by GFFread version 0.12.3 (Pertea & Pertea, 2020) using the settings --adj-stop -J --sort-alpha -E --keep-genes as the pred_gff passed to MAKER. We functionally annotated the MAKER filtered genes using proteins with a combination of blastp searches against UniProt/Sprot release 2020_05 implemented with Diamond version 2.0.4 (Buchfink et al., 2015) using the settings ultra-sensitive, evalue 1e-6 and max-target-seqs 1. The resulting annotations were reformatted with GAG (http://genomeannotation.github.io/GAG/; Geib et al., 2018) and Annie (http://genomeannotation.github.io/annie/).

In addition, we annotated the noncoding RNA genes, including transfer RNA (tRNA) genes and ribosomal RNA (rRNA) genes, within the genome assembly. tRNAscan-SE version 2.0 (Lowe & Eddy, 1997) with Eukaryotic parameters were used to predict the tRNA genes. The rRNA genes were annotated using RNAmmer version 1.2 (Lagesen et al., 2007) with default parameters.

2.7 | Phylogenetic analysis

To determine the phylogenetic relationship among the dipterous flies, we retrieved protein sets of *Lucilia cuprina* (ASM118794v1), *Musca domestica* (MdomA1), *Stomoxys calcitrans* (ScalU1), *Glossina morsitans* (GmorY1), *Drosophila melanogaster* (BDGP6.32), *Mayetiola destructor* (Mdes_1.0), *Aedes aegypti* (AaegL5), *Anopheles gambiae* (AgamP4) from Ensembl Metazoa and *Sarcophaga bullata* (GCA_005959815.1) from NCBI. For genes with more than two transcripts within genomes, we only kept the protein sequence of the longest transcript. Then, gene families were clustered with OrthoFinder version 2.5.1 with the settings: -M msa -S blast -A mafft -T fasttree (Emms & Kelly, 2015, 2019), which specifies multiple sequence alignments (–M) for the gene tree inference. The protein sequences of the resulting single-copy genes were aligned using Mafft version 7.475 (Katoh & Standley, 2013) with default parameters, followed by trimming with gBlocks to remove gaps (Talavera & Castresana, 2007). After trimming, we used SeqKit version 0.10.0 (Shen et al., 2016) to concatenate the trimmed protein sequences of single-copy orthologous of each species into one super gene. We used ProtTest version 3.4.2 (Darriba et al., 2011) to determine the optimal amino acid substitution model. Subsequently, the phylogenetic trees were inferred using RAxML version 8.2.12 (Stamatakis, 2014) using the PROTGAMMALG substitution model with 1000 bootstrap replicates.

To estimate divergence times among species, the MCMCTree program of Paml version 4.9 (Yang, 2007) was employed and five calibration points were obtained from the TimeTree database (http://www.timetree.org/), including *M. domestica–S. calcitrans* (27–37 million years ago [Ma]), *M. domestica–L. cuprina* (47–71 Ma), *M. domestica–G. morsitans* (48–74 Ma), *M. domestica–D. melanogaster* (107–172 Ma), and *A. gambiae–A. aegypti* (52–147 Ma).

2.8 | Analysis of parasitism-related genes

To identify genes associated with parasitism, we selected the protein sets with the longest transcript of three myiasis-causing flies, including *C. hominivorax*, *L. cuprina* and *W. magnifica*, and *D. melanogaster*, which feeds on rotting fruit, and clustered their gene families using OrthoFinder version 2.5.1 with the same settings as above. Then, the genes in the resulting gene families shared by three myiasis-causing flies and absent in *D. melanogaster* were further annotated and analysed.

2.9 | Gene family expansion and contraction

Based on the clustering of gene families generated by OrthoFinder and the phylogenetic relationship with divergence times determined by RAxML and MCMCTree, we used Café version 4.1 (De Bie et al., 2006) to analyse the expansion and contraction of gene families, which uses a birth and death process to model gene gain and loss over phylogenetic distance. The resulting expanded genes were extracted for gene ontology (GO) enrichment analysis in OmicShare

tools (https://www.omicshare.com/tools) using a false discovery rate (FDR) < 0.05 for multiple test correction.

2.10 | Positive selection analysis

Including all single-copy orthologues of the 10 dipterous flies previously inferred by OrthoFinder and their corresponding coding sequences (CDS), we used ParaAT version 1.0 (Zhang et al., 2012) to align single-copy orthologues and then back-translate the multiple protein sequence alignment into a codon alignment with the settings: -m mafft -g -t. Next, the codeml program in the Paml package version 4.9 (Yang, 2007) was implemented with the alignment results as inputs using the branch-site model with the *W. magnifica* branch as foreground and the remaining dipterous fly branches as background. Then, we compared the alternative model (model = 2, NSsites = 2, fix_omega = 0) to the null model (model = 2, NSsites = 2, fix_omega = 1 and omega = 1) using a likelihood ratio test (LRT) calculated with a Chi-square distribution ($p < .05$; one degree of freedom [$df = 1$]). We corrected for multiple testing using FDR < 0.05 and retained only genes that contained amino-acid sites of positive selection ≥1 as final positive selection candidates.

3 | RESULTS

3.1 | Genome assembly and assessment of *W. magnifica*

Since *W. magnifica* cannot be reared in the laboratory, inbred lines cannot be obtained. Taking advantage of the protocol previously described in the method section, which requires a minimum amount of DNA of only >400 ng, we successfully prepared 10 kb Pacbio libraries for sequencing on Pacbio Sequel II System using 1.215 μg of high-quality genomic DNA extracted from a single female adult *W. magnifica*.

A total of approximately 408 Gb of raw data composed of 43,228,999 reads with 9447 bp average sequence length were produced on a single SMRT Cell 8 M. After converting, a total of approximately 23 Gb of HiFi sequences comprised of 2,197,069 HiFi reads with 10,681 bp average length were obtained, which is approximately 30x coverage based on the genome assembly of *W. magnifica* (Table 1).

Subsequently, HiFi sequences were assembled using the IPA program. As a result, we obtained a set of 753.99 Mb of primary contigs used as the assembled genome and a set of 647.62 Mb of alternative contigs. The assembled genome has a scaffold N50 of 5.00 Mb, the longest scaffold length of 14.66 Mb, a scaffold number of 543 and a proportion of the bases guanine and cytosine of 32.82% (Table 4). Compared to other dipterous fly genomes, *W. magnifica* has a similar genome size to *M. domestica* but is more than five times larger than the genome of *D. melanogaster*. It is also larger than the genome of *S. bullata*, which belongs to the same family Sarcophagidae, and *C. hominivorax*, which has a similar way of invading livestock to *W. magnifica* (Table 2).

We evaluated the quality of the genome assembly using BUSCO program to search against 3285 conserved single-copy genes (diptera_odb10). This analysis indicated that 98.8% complete BUSCO genes (3245 genes), including 98.2% complete and single-copy (3226 genes) and 0.6% complete and duplicated (19 genes), and 0.6% fragmented BUSCO genes (19 genes) could be captured, with only 0.6% missing (21 genes). The BUSCO results were comparable to that of nine publicly available dipterous fly genomes (Figure 1a, Table S2). In addition, we mapped the RNA-seq data of *W. magnifica* at different developmental larvae stages towards the genome assembly, which was subsequently used in the BRAKER program to aid in the gene structure annotation of the genome, resulting in a mapping rate of 93.62%. Taken together, these results suggested that the genome is complete and accurate.

3.2 | Annotation of the de novo assembled genome of *W. magnifica*

We searched repeat elements in the assembled genome of *W. magnifica* using Repeatmasker with a species-specific repeat library generated by RepeatModeler. Overall, identified repetitive sequences accounted for 59.71% (450 Mb) of the total assembled genome, consisting of 56.76% interspersed, 2.48% simple repeat sequences and 0.4% low complexity. Among the interspersed repeats, the most abundant were unassigned sequences (23.37% of the assembled genome), followed by DNA transposons (16.62%), LINEs (12.92%), and

TABLE 1 Sequencing data statistics of *Wohlfahrtia magnifica*

Parameter	Raw data	Converted HiFi data
Total sequence length (bp)	408,396,865,291	23,467,199,316
Total sequence number	43,228,999	2,197,069
Average sequence length (bp)	9447	10,681
GC content (%)	33.79857849	32.7324404
Max sequence length (bp)	468,940	38,618
N20 (bp)	14,919	15,097
N50 (bp)	10,949	11,580
N90 (bp)	6089	7338

TABLE 2 Comparison of summary statistics of genome assembly between *Wohlfahrtia magnifica* and other dipterous flies

Species	*W. magnifica*	*M. domestica*	*D. melanogaster*	*S. bullata*	*C. hominivorax*
Genome size	753.99 Mb	750.40 Mb	143.73 Mb	476.29 Mb	534.08 Mb
Number of scaffolds	543	20,487	1870	42,093	3663
Scaffold N50	5.00 Mb	226.57 kb	25.29 Mb	55.53 kb	616.42 kb
Reference	In the study	GCA_000371365.1	GCA_000001215.4	GCA_005959815.1	Scott et al. (2020) and GCA_004302925.1

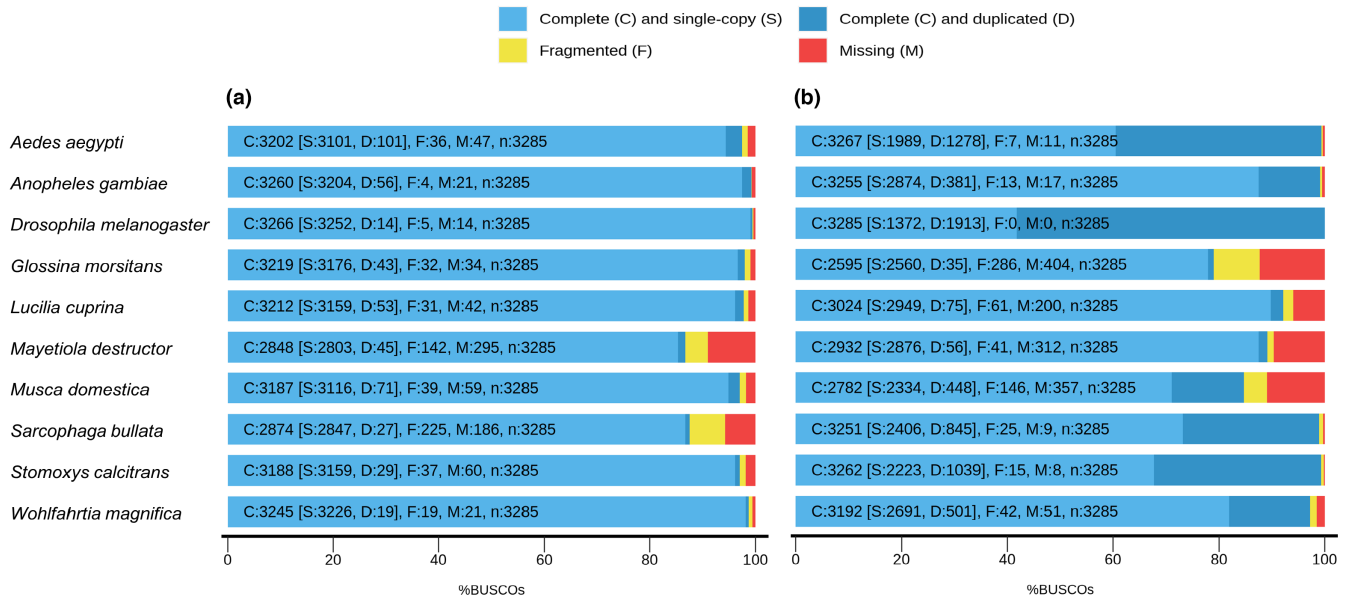

FIGURE 1 BUSCO analysis between *Wohlfahrtia magnifica* and other nine dipterous flies based on diptera_odb10 gene set. (a) BUSCO assessment results of genomes. (b) BUSCO assessment results of gene sets. C, complete BUSCO genes; S, single-copy BUSCO genes; D, duplicated BUSCO genes; F, fragmented BUSCO genes; M, missing BUSCO genes; *n*, total number of BUSCO genes

TABLE 3 Repeat element statistics of *Wohlfahrtia magnifica* assembled genome

Repeat element	Numbers	Bases	% of genome
SINEs	0	0	0.00
LINEs	313,495	97,444,031	12.92
LTR	98,752	29,004,031	3.85
DNA	473,914	125,334,368	16.62
Unclassified	976,699	176,211,233	23.37
Simple repeats	279,951	18,690,856	2.48
Low complexity	61,088	2,981,425	0.40
Total	/	450,243,625	59.71

LTR elements (3.85%) (Table 3). In addition, we compared the proportion of the genome occupied by repetitive sequences between *W. magnifica* and the other three dipterous flies. The results showed that the repetitive sequences in the genome of *W. magnifica* were similar to *L. cuprina* (57.82%), but significantly higher than those of *C. hominivorax* (45.22%) and *S. bullata* (31.15%) (Table S3).

We employed the BRAKER2 pipeline followed by MAKER2 to predict the gene model with the aid of transcriptomic data and protein data (Arthropoda v100_odb10). As a result, 16,718 protein-encoding genes and 20,017 transcripts were identified in the assembled genome of *W. magnifica*. The longest, shortest gene length and mean gene length as well as mean length for exon, intron and coding sequence are given in Table 4. Of the genes, 64.98% could be functionally annotated in the UniProt/Sprot database. In addition, the gene set was assessed by the BUSCO program with proteins mode, indicating that 97.2% complete conserved single copy genes (diptera_odb10) could be identified, whereas only 1.5% were assigned as missing. This is consistent with other Diptera genomes, suggesting the gene annotation of our de novo assembled genome is of comparable completeness (Figure 1b, Table S2).

Moreover, we identified and annotated noncoding RNA genes in the genome of *W. magnifica*, showing that 576 tRNAs (excluding tRNAs identified as pseudogenes) and 53 rRNAs were identified.

3.3 | Phylogenetic analysis of *W. magnifica*

We searched the orthologues among predicted proteins of *W. magnifica* and those derived from the other nine dipterous flies using OrthoFinder program. A total of 149,614 genes were recovered and 135,230 were grouped into 14,424 orthogroups. The remaining genes were clustered into 14,384 unassigned species-specific orthogroups, of which each consisted of only one gene. Of 14,424

orthogroups, 2045 orthogroups were found in the gene sets of all 10 flies in single copy form. In addition, we detected 1736 genes present only in the *W. magnifica* genome, including 943 multiple-copy and 793 single-copy genes (Figure 2a, Table S4). We also found 7306 orthogroups present in all four dipterous flies, including *W. magnifica* and its closest evolutionary relatives, *S. bullata*, and *L. cuprina*, as well as *D. melanogaster* (Figure 2b).

TABLE 4 Protein-coding gene annotation statistics of *Wohlfahrtia magnifica* assembled genome

Parameter	
Genome size	753.99 Mb
Number of contigs	543
Contig N50	5.00 Mb
GC content	32.82%
Max scaffold length	14.66 Mb
Number of scaffolds >50 KB	541
Number of genes	16,718
Number of mRNAs	20,017
Number of exons	87,424
Number of introns	67,407
Number of CDS	20,017
Shortest gene	147 bp
Longest gene	394,287 bp
Mean gene length	9789 bp
Mean mRNA length	11,538 bp
Mean exon length	347 bp
Mean intron length	2978 bp
Mean CDS length	1515 bp
Mean exons per mRNA	4
Mean introns per mRNA	3
% of genome covered by genes	21.7%
% of genome covered by CDS	4.0%

Using the above-obtained single-copy orthologues, we performed the phylogenetic reconstruction. Our results suggested that *W. magnifica* is the closest phylogenetic relative to *S. bullata*, as these two dipterous flies belong to the family Sarcophagidae, followed by *L. cuprina*, while being most distant from *A. aegypti* and *A. gambiae* (Figure 3a). As we expected, this result is consistent with other dipteran phylogenetic trees (Martinson et al., 2019; Scott et al., 2020). Furthermore, divergence time estimation revealed that the common ancestors of *W. magnifica* and *S. bullata* split from *L. cuprina* approximately 30.51 Ma, while *W. magnifica* and *S. bullata* divergence time were dated to 19.81 Ma (Figure 3a).

3.4 | Analysis of parasitism-related genes

We found 885 gene families shared by three myiasis-causing flies, but no orthologue in *D. melanogaster*, composed of 1707 genes in *W. magnifica*, 1211 genes in *L. cuprina* and 2502 genes in *C. hominivorax*, which may play an important role in parasitism. Further analysis of 1707 genes in *W. magnifica* revealed that 1548 (90.69%) of the genes could be annotated in the NR database. Of 1548 genes, 973 genes were annotated as experimentally uncharacterized genes. Analysis of the remaining 575 genes with specific functional annotations found some genes of interest in parasitism, such as olfactory-related genes for site search for laying the larvae, insecticide resistance-related genes and protease for digestion of host tissues (Data S1).

3.5 | Gene family expansion and contraction of *W. magnifica*

The expansion and contraction of a species' gene family is often associated with its adaptive evolution. In the genome of *W. magnifica*,

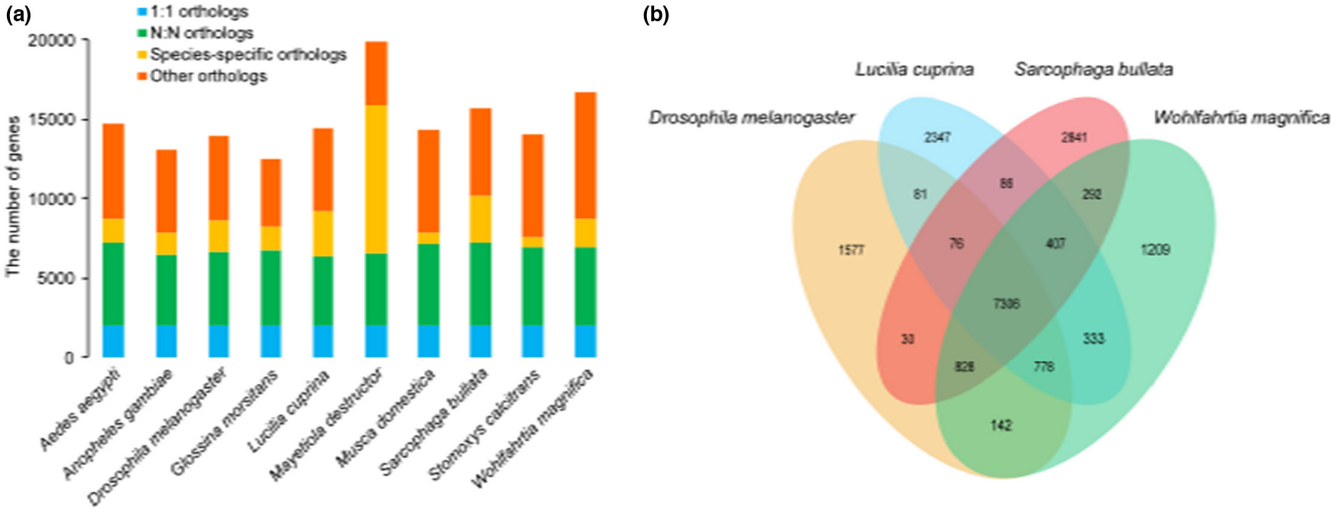

FIGURE 2 Gene family analysis of *Wohlfahrtia magnifica* and other dipterous flies. (a) Distribution of genes in different species. 1:1 orthologues: Single-copy orthologues; N:N orthologues: Multiple-copy orthologues; species-specific orthologues: Present in specific species; other orthologues: The remaining orthologues. (b) Numbers of orthogroups shared or unique in *W. magnifica*, *S. bullata*, *L. cuprina* and *D. melanogaster*

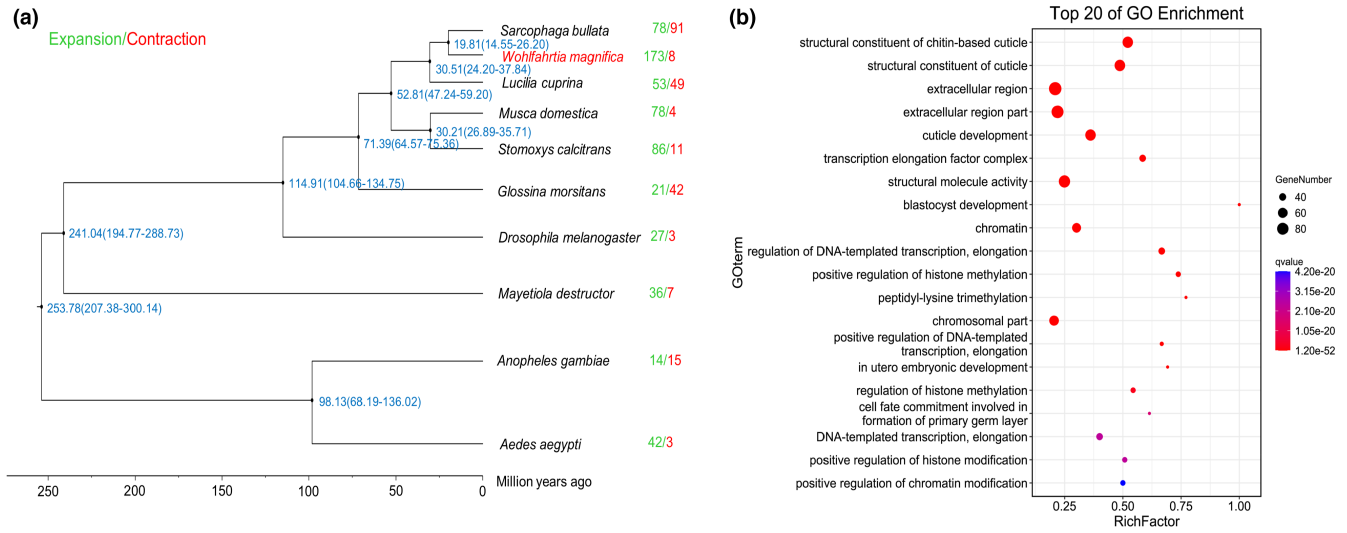

FIGURE 3 Comparative analysis of the genome of *Wohlfahrtia magnifica*. (a) Phylogenetic relationship, the estimated divergence time (Ma) and gene family of expansion and contraction (the red numbers reflect contracted gene families, whereas the green numbers show expanded gene families.) of *W. magnifica* and other nine dipterous flies. (b) Top 20 of GO enrichment results of 173 expanded gene families of *W. magnifica*

we identified 173 expanded gene families, comprising 2008 genes, and eight contracted families. Subsequently, a GO enrichment study of expanded gene families revealed 24 GO terms for cellular component categories, 10 GO terms for molecular function categories and 117 GO terms for the biological process categories (Figure 3b). We deemed the following GO terms of special interest as these terms likely help to improve the understanding of adaptation of *W. magnifica* to parasitic life-style: structural constituent of chitin-based cuticle (GO:0005214, *p*.adjust = $1.17e^{-52}$); structural constituent of cuticle (GO:0042302, *p*.adjust = $6.96e^{-51}$); cuticle development (GO:0042335, *p*.adjust = $1.10e^{-35}$); response to insecticide (GO:0017085, *p*.adjust = $4.23e^{-21}$); response to toxic substance (GO:0009636, *p*.adjust = $8.57e^{-15}$); response to bacterium (GO:0009617, *p*.adjust = $6.69e^{-12}$); defence response to bacterium (GO:0042742, *p*.adjust = $1.91e^{-09}$) and instar larval development (GO:0002168, *p*.adjust = $1.27e^{-07}$) (Figure 3b, Table S5). In addition, we also found that many GO terms are involved in transcription and chromatin modification, such as transcription elongation factor complex (GO:0008023, *p*.adjust = $6.44e^{-34}$); regulation of DNA-templated transcription, elongation (GO:0032784, *p*.adjust = $9.12e^{-31}$); positive regulation of chromatin modification (GO:1903310, *p*.adjust = $4.21e^{-20}$) and positive regulation of histone modification (GO:0031058, *p*.adjust = $2.43e^{-20}$).

3.6 | Genes under positive selection in *W. magnifica*

To determine genes potentially under positive selection in *W. magnifica*, 2045 single-copy orthologues between *W. magnifica* and the other nine dipterous flies were multiply protein-aligned and back-translated. After this, we discarded 61 orthologues due to poorly aligned positions and retained 1984 orthologues, which were subjected to the branch-site analysis. In the *W. magnifica* branch, 45 genes were identified as likely under positive selection. However, the annotated function of *D. melanogaster* orthologues of 13 of these genes was unknown. We investigated the functions of the remaining 32 genes, showing that the function of these genes spans a wide range of areas. Of these genes, the following genes are of particular interest: development-related genes, including neuronal development (*Plex A and D*), muscle development (*Wnt2* and *Kon-tiki*) and eye development (*Myt1*), Ca(2+) regulation (*RyR*), regulation of metabolism (*Wdr24*) and melanic pigmentation (*yellow-k*) (Table S6).

4 | DISCUSSION

4.1 | De novo genome assembly from a single *W. magnifica* female using a low-input DNA workflow

Genomic resources are essential for a thorough understanding of the molecular biology of an organism and the evolution of a species. In addition, they can provide insights into specific mechanisms of interest, for example in relation to environmental adaption, production traits in livestock, or for designing new control strategies of invasive species, for example, the spotted flesh fly (*W. magnifica*). Considering the small size of a single individual and the difficulty of obtaining inbreed of *W. magnifica*, we exploit the advantages of low DNA input and highly accurate HiFi sequence when using the low-input DNA workflow for HiFi library preparation and sequencing on Pacbio sequel II with CCS mode. As a result, the de novo assembled genome consists of only 543 scaffolds with a scaffold N50 of 5.00 Mb. We used BUSCO and other metrics to assess the quality of the assembly and the results suggested that the de novo genome is of high quality, only with a few genes missing. In addition,

our assembled genome is comparable to other genomes sequenced using various strategies such as *L. cuprina* (Anstead et al., 2015), *C. hominivorax* (Scott et al., 2020) or *S. peregrina* (Ren et al., 2021), both in terms of N50 and BUSCO results. Therefore, the strategy used in this study can serve as a reference sequencing approach for some dipterous flies that are small and not easy to rear in the lab. This also can save time for flies which can be reared in the lab but are challenging to inbreed. Importantly, this strategy will facilitate further insect sequencing projects like the 5000 Insect Genome Project (i5k) (i5K Consortium, 2013). Although our currently assembled genome has a very high quality, which can meet the requirements for genomic applications in *W. magnifica's* pest control, such as the development of Cas9-based homing gene drives strains (Hammond et al., 2016) and transgenic sexing strains (Concha et al., 2020; Li et al., 2014), we suggest using Hi-C technology to capture the organizational structure of chromatin in three dimensions. This will enhance further analysis, such as identification of promoter-enhancer interactions for gene regulation studies and detection of structural rearrangements.

4.2 | Parasitism-related genes, gene families of expansion and positively selected genes related to adaptation and evolution in *W. magnifica*

Myiasis is known as a disease of living vertebrates invaded by dipteran larvae, which is of great medical and veterinary importance, as it affects not only wild and domestic animals but also humans in developed and developing countries worldwide (Zumpt, 1965). Based on the dependence degree to host, myiasis can be classified into three types: accidental, facultative or obligatory myiasis (Scholl et al., 2019). In this study, to identify genes associated with parasitism we selected three obligatory myiasis-causing flies, *C. hominivorax*, *L. cuprina* and *W. magnifica*, and one nonmyiasis-causing fly, *D. melanogaster*. A number of parasitism-related genes were identified, such as olfactory-related genes, proteases and some insecticide-resistant genes. Usually, myiasis-causing flies use their olfactory system to detect the odour from the host's open wound or genitalia while looking for a site to lay their eggs or larvae. Thus, these olfactory genes, such as odorant receptor and odorant-binding protein, may be involved in the behaviour of the search for egg-laying sites. Once these larvae have been oviposited into the host, these proteases help the larvae digest the host's tissues into small molecule peptides and amino acids for development. In addition, we also found several insecticide-resistant genes in the gene sets, especially cytochrome P450 (CYP450), which may confer the resistance of myiasis-causing flies to insecticide. For example, *L. cuprina* has developed resistance to organochlorines (e.g., dieldrin/aldrin), organophosphates (OP) (e.g., diazinon), carbamate (e.g., butacarb) and others (Sandeman et al., 2014). Although we found many genes of interest in this gene set, there are still many genes that are defined as hypothetical proteins with little to no experimental evidence for their function or being characterized by a low identity to proteins with known function. However, these genes may play an essential role in parasitism of myiasis-causing flies.

We also investigated which gene families expanded and which GO categories were enriched by these expanding families. This may enhance our understanding of the adaption of *W. magnifica* to its parasitic lifestyle and may help identify potential strategies for pest control. For example, from late April or early May to mid-October in China, as *W. magnifica* infects its hosts, large quantities of insecticides are used to kill the larvae of *W. magnifica*. This likely leads to adaptation in *W. magnifica's* response to insecticide, as the GO terms "response to insecticide" and "response to toxic substance" are enriched. The body temperature of camels varies considerably, from 40°C during the day to 34°C at night (Schmidt-Nielsen et al., 1956). Consequently, this may induce thermal stress in *W. magnifica* larvae. This might explain the expansion of gene families with GO terms such as "response to temperature stimulus", "cellular response to heat" and "response to heat". Additionally, bacteria grow on the wound of the infected host and the host's immune response may affect the larval environment. This may in part explain the GO enrichment results also identified a number of expanded gene families with genes enriched for infection response-related terms, for example, "response to bacterium", "defence response to bacterium", "defence response to other organism", "mucosal immune response" and "organ or tissue specific immune response". The cuticle, acting as a barrier between living tissues and the surrounding atmosphere, is a multilayered structure, which has various functions, for example, the determination of the shape and appearance, insecticide resistance and constituting a physical barrier to prevent pathogen entry (Andersen, 1979; Balabanidou et al., 2018; Moussian, 2010). The current study found that gene families that expanded most were associated with several GO categories (according to *p*-value) linked to cuticle development. These results indicate that the cuticle may play a very important role in the adaptation of *W. magnifica* to its parasitic lifestyle. Interestingly, we also found that many of the GO categories are involved in transcription and chromatin modification, which is responsible for the regulation of gene expression, suggesting there might have been some major changes in gene expression during the evolution of *W. magnifica's* parasitic lifestyle.

In this study, we obtained up to 45 positively selected genes with diverse functions and molecular processes. These genes subjected to positive selection are likely to contribute to *W. magnifica's* evolution and adaptation. For *W. magnifica*, the developmental stage of the embryo and larvae is inside the female fly and inside the open wound or genitalia of the host, respectively. Of positively selected genes, we found several genes associated with development, including neuronal development (*Plex A and D*) (Junqueira Alves et al., 2019; Overton et al., 2002; Soriano & Russell, 1998), muscle development (*Wnt2* and *Kon-tiki*) (Estrada et al., 2007; Schnorrer et al., 2007) and eye development (*Myt1*) (Price et al., 2002). These genes may contribute to the adaptation of *W. magnifica's* embryo and larvae to in vivo development. Insects are able to find the location of food, mates, and egg-laying sites with the help of their olfactory systems (De Bruyne

& Baker, 2008; He et al., 2019). In *Drosophila*, interfering with *RyR* expression resulted in a defective olfactory behaviour in flies (Murmu et al., 2010). Therefore, this gene might have an essential role in *W. magnifica's* response to several odours from the host's wounds or genitalia during the search for egg-laying sites. *Wdr24*, a component of a multiprotein GATOR2 complex, is a critical part of the cellular metabolism, such as nutrients in different species, including *Drosophila*. During the development of *W. magnifica*'s larvae, the nutrients are mainly derived from the host tissues. Perhaps *Wdr24* plays an important role in tissue metabolism (Cai et al., 2016; Kim, Nam, et al., 2019; Kim, Paggi, et al., 2019). *Drosophila's* melanic pigmentation in the wings, abdomen and thorax is now recognized to be related to the *yellow* locus (Ferguson et al., 2011). Compared to other myiasis-causing flies, the abdomen of *W. magnifica* has distinguishing dark-coloured spots, which might be associated with *yellow-k*.

4.3 | Potential applications for the control of *W. magnifica*

Although the larvae of *W. magnifica* parasitize several warm-blooded vertebrates, in China, its primary host is camels. During the summer months when *W. magnifica's* populations are high, grazing Bactrian camels are present across the Gobi Desert or grassland, and therefore are not frequently inspected, resulting in infected camels not being treated in a timely manner and aggravating the infection. The severe infection in this condition poses a threat to important animal welfare and health of Bactrian camels, and induces especially reproduction problems, such as abortion. Unfortunately, the infection also affects the wild camel (*Camelus ferus*), which is listed as Critically Endangered by the International Union for the Conservation of Nature (IUCN). It is estimated that there are approximately 1000 individuals left, around 600 in the Gobi Desert in northwest China and probably only 450 at the Mongolian side. (https://www.wildcamels.com/). The threat to the wild camel can be especially devastating because *W. magnifica* is unmanageable in wildlife populations. Therefore, similar to Bactrian camel, vital research on the control of *W. magnifica* is important for the conservation of this critically endangered wild camel.

In contrast to *W. magnifica*, *C. hominivorax* was successfully eradicated from the USA and Central America by the application of the sterile insect technique (SIT) (Wyss, 2006). With regards to *W. magnifica*, the assembled genome could pave the path for the identification of reproduction-related genes, which might contribute towards the development of further SIT (Baumhover et al., 1955; Knipling, 1955). However, for SIT or other methods of genetic control of *W. magnifica*, due to the current high mortality rate of rearing in the laboratory, this would require significant progress in methods for rearing *W. magnifica* on artificial diet. For the prevention and control of *W. magnifica* in camels and other livestock, vaccines might be an effective strategy. On the basis of the de novo assembled genome of *W. magnifica*, a great range of candidate vaccine antigens might be identified, and effective antigens likely involved in inducing a protective immune response of the infected host against *W. magnifica* at larval stages could be selected for the development of subunit vaccines in the future. So far, chemical control methods against *W. magnifica* infections dominate in China. However, the excessive use of insecticides can lead to insecticide resistance, as supported by our results from the analysis of expanded gene families. Genome-guided identification is a comprehensive and promising strategy to screen new drug targets and discover new drugs (Olsen & Faergeman, 2012). The approach aims to identify candidate genes or gene products that can be inactivated by insecticides, without harming to the host animal. For investigation of gene functions and insecticidal target discovery, the RNA interference (RNAi), combined with the resulting phenotype, is an effective approach, because it has been a great success with many insects (Hu et al., 2016; Riga et al., 2020). For the evaluation of gene functions on a genome-wide scale in *W. magnifica*, RNAi is not for routine use. In this case, essential single-copy genes of *W. magnifica* can be predicted using functional genomic data (e.g., lethality) available for orthologues in *D. melanogaster*, for which potential insecticidal targets have already been identified (Anstead et al., 2015; Olsen & Faergeman, 2012). In addition, genome-wide identification of complete chemosensory genes could likely be beneficial for suppressing the *W. magnifica* population and monitoring its behaviour by trapping flies by odours. Clearly, the de novo assembled genome has enabled us to enter an exciting era in which the door to the development or improvement of novel genetic, immunological and chemical control strategies for *W. magnifica* is opened.

In conclusion, we successfully assembled de novo the genome of *W. magnifica* using only one female adult individual. This assembled genome is 753.99 Mb in size with a scaffold N50 length of 5.00 Mb and 59.71% repeat elements. The RNA-seq mapping rate and BUSCO scores indicate that the assembled genome is complete (93.62% overall RNA-seq alignment rate and 98.8% complete BUSCOs found). In addition, 16,718 genes and 20,017 mRNA were predicted in the assembled genome; of these, 64.98% of genes can be functionally annotated in the UniProt/Sprot database. Phylogenetic analysis revealed that *W. magnifica* is most closely related to *S. bullata*, followed by *L. cuprina*. GO enrichment analysis showed that many of the expanded gene families contained genes annotated for immunity, insecticide-resistance mechanisms, heat stress response and cuticle development, while positively selected genes displayed diverse functions. Clearly, the availability of the current *W. magnifica* genome resource lays a solid foundation for being able to address key biological questions and to facilitate the development of new prevention and control methods of this mammal's pest in the future.

AUTHOR CONTRIBUTIONS

Zhipeng Jia and Pamela A. Burger conceived the project and received funding. Zhipeng Jia performed the genome annotation and comparative genomic analysis and wrote the first draft of the manuscript. Surong Hasi and Claus Vogl contributed new reagents, samples or analytical support. Pamela A. Burger, Surong Hasi and Claus

Vogl supervised the project. All authors provided valuable advice, reviewed and approved the final manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge J.-P. Elbers for excellent bioinformatic support and Ming Liang for technical help. We are very grateful to the camel owner and members of Professor Surong Hasi Laboratory for their assistance in the sample collection of *W. magnifica* for scientific purposes. The work was supported by the China Scholarship Council (201909150004) granted to Z.J. and the Austrian Science Fund (FWF) project (P29623-B25) to P.B.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The genome assembly of *W. magnifica* was deposited as a BioProject under accession number PRJNA778059. The PacBio HiFi sequence reads are deposited at NCBI under accession number: SRR16848117. The transcriptome data have been deposited in SRA, including 3 first stage larvae (SRR18178228, SRR18178229, SRR18178230), 3 second stage larvae (SRR18178225, SRR18178226, SRR18178227) and 3 adult flies (SRR18178222, SRR18178223, SRR18178224). In addition, the assembly and annotation of the W. magnifica genome are also available on Dryad (https://doi.org/10.5061/dryad.qfttdz0j8).

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.org/10.5061/dryad.qfttdz0j8.

BENEFITS-SHARING STATEMENT

A research collaboration was developed with scientists from the countries providing genetic samples. All collaborators are included as coauthors. Benefits from this research accrue from the sharing of our data and results on public databases as described above.

ORCID

Zhipeng Jia https://orcid.org/0000-0001-6524-2957
Pamela A. Burger https://orcid.org/0000-0002-6941-0257

REFERENCES

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., … Venter, J. C. (2000). The genome sequence of *Drosophila* melanogaster. *Science*, *287*(5461), 2185–2195. https://doi.org/10.1126/science.287.5461.2185

Andersen, S. O. (1979). Biochemistry of insect cuticle. *Annual Review of Entomology*, *24*(1), 29–61. https://doi.org/10.1146/annurev.en.24.010179.000333

Anstead, C. A., Korhonen, P. K., Young, N. D., Hall, R. S., Jex, A. R., Murali, S. C., Hughes, D. S., Lee, S. F., Perry, T., Stroehlein, A. J., Ansell, B. R., Breugelmans, B., Hofmann, A., Qu, J., Dugan, S., Lee, S. L., Chao, H., Dinh, H., Han, Y., … Gasser, R. B. (2015). Lucilia cuprina genome unlocks parasitic fly biology to underpin future interventions. *Nature Communications*, *6*(1), 1–11. https://doi.org/10.1038/ncomms8344

Balabanidou, V., Grigoraki, L., & Vontas, J. (2018). Insect cuticle: A critical determinant of insecticide resistance. *Current Opinion in Insect Science*, *27*, 68–74. https://doi.org/10.1016/j.cois.2018.03.001

Baumhover, A. H., Graham, A. J., Bitter, B. A., Hopkins, D. E., New, W. D., Dudley, F. H., & Bushland, R. C. (1955). Screw-worm control through release of sterilized flies. *Journal of Economic Entomology*, *48*(4), 462–466. https://doi.org/10.1093/jee/48.4.462

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. https://doi.org/10.1038/nmeth.3176

Cai, W., Wei, Y., Jarnik, M., Reich, J., & Lilly, M. A. (2016). The GATOR2 component Wdr24 regulates TORC1 activity and lysosome function. *PLoS Genetics*, *12*(5), e1006036. https://doi.org/10.1371/journal.pgen.1006036

Campbell, M. S., Law, M., Holt, C., Stein, J. C., Moghe, G. D., Hufnagel, D. E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C. J., Ware, D., Shiu, S. H., Childs, K. L., Sun, Y., Jiang, N., & Yandell, M. (2014). MAKER-P: A tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiology*, *164*(2), 513–524. https://doi.org/10.1104/pp.113.230144

Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., & Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188–196. https://doi.org/10.1101/gr.6743907

ÇiftÇio, N., Altintaş, K., & Haberal, M. (1996). A case of human orotracheal myiasis caused by Wohlfahrtia magnifica. *Parasitology Research*, *83*(1), 34–36. https://doi.org/10.1007/s004360050203

Concha, C., Yan, Y., Arp, A., Quilarque, E., Sagel, A., de León, A. P., WO, M. M., Skoda, S., & Scott, M. J. (2020). An early female lethal system of the New World screwworm, Cochliomyia hominivorax, for biotechnology-enhanced SIT. *BMC Genetics*, *21*(2), 1–12. https://doi.org/10.1186/s12863-020-00948-x

Cruz, M. D. S., Robles, M. C. V., & Thomas, G. (1996). In vivo rearing and development of Wohlfahrtia magnifica (Diptera: Sarcophagidae). *Journal of Medical Entomology*, *33*(4), 586–591. https://doi.org/10.1093/jmedent/33.4.586

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics*, *27*(8), 1164–1165. https://doi.org/10.1093/bioinformatics/btr088

De Bie, T., Cristianini, N., Demuth, J. P., & Hahn, M. W. (2006). CAFE: A computational tool for the study of gene family evolution. *Bioinformatics*, *22*(10), 1269–1271. https://doi.org/10.1093/bioinformatics/btl097

De Bruyne, M., & Baker, T. C. (2008). Odor detection in insects: Volatile codes. *Journal of Chemical Ecology*, *34*(7), 882–897. https://doi.org/10.1007/s10886-008-9485-4

Dehghani, R., Zarghi, I., & Sayyedi, H. R. (2014). Genital myiasis of a sheep by Wohlfahrtia magnifica, in Ghamsar, Kashan, Iran. *Bangladesh Journal of Medical Science*, *13*(3), 332–335. https://doi.org/10.3329/bjms.v13i3.15451

Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 1–14. https://doi.org/10.1186/s13059-015-0721-2

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 1–14. https://doi.org/10.1186/s13059-019-1832-y

Estrada, B., Gisselbrecht, S. S., & Michelson, A. M. (2007). The transmembrane protein Perdido interacts with grip and integrins to

mediate myotube projection and attachment in the drosophila embryo. *Development*, *134*(24), 4469–4478. https://doi.org/10.1242/dev.014027

Farkas, R., Hall, M. J., Daniel, M., & Börzsönyi, L. (1996). Efficacy of ivermectin and moxidectin injection against larvae of Wohlfahrtia magnifica (Diptera: Sarcophagidae) in sheep. *Parasitology Research*, *82*(1), 82–86. https://doi.org/10.1007/s004360050073

Farkas, R., Hall, M. J., & Kelemen, F. (1997). Wound myiasis of sheep in Hungary. *Veterinary Parasitology*, *69*(1–2), 133–144. https://doi.org/10.1016/s0304-4017(96)01110-7

Farkas, R., Hell, E., Hall, M. J. R., & Gyurkovszky, M. (2005). In vitro rearing of the screwworm fly Wohlfahrtia magnifica. *Medical and Veterinary Entomology*, *19*(1), 22–26. https://doi.org/10.1111/j.0269-283x.2005.00529.x

Farkas, R., & Képes, G. Y. (2001). Traumatic myiasis of horses caused by Wohlfahrtia magnifica. *Acta Veterinaria Hungarica*, *49*(3), 311–318. https://doi.org/10.1556/004.49.2001.3.8

Ferguson, L. C., Green, J., Surridge, A., & Jiggins, C. D. (2011). Evolution of the insect yellow gene family. *Molecular Biology and Evolution*, *28*(1), 257–272. https://doi.org/10.1093/molbev/msq192

Gaglio, G., Brianti, E., Abbene, S., & Giannetto, S. (2011). Genital myiasis by Wohlfahrtia magnifica (Diptera, Sarcophagidae) in Sicily (Italy). *Parasitology Research*, *109*(5), 1471–1474. https://doi.org/10.1007/s00436-011-2431-3

Geib, S. M., Hall, B., Derego, T., Bremer, F. T., Cannoles, K., & Sim, S. B. (2018). Genome annotation generator: A simple tool for generating and correcting WGS annotation tables for NCBI submission. *GigaScience*, *7*(4), 1–5. https://doi.org/10.1093/gigascience/giy018

Giangaspero, A., Traversa, D., Trentini, R., Scala, A., & Otranto, D. (2011). Traumatic myiasis by Wohlfahrtia magnifica in Italy. *Veterinary Parasitology*, *175*(1–2), 109–112.

Hall, M. J. (1997). Traumatic myiasis of sheep in Europe: A review. *Parassitologia*, *39*(4), 409-13.

Hall, M. J. R., & Farkas, R. (2000). Traumatic myiasis of humans and animals. In L. Papp & B. Darvas (Eds.), *Contributions to a manual of Palaearctic Diptera* (pp. 751–768). Science Herald.

Hall, M. J. R., Testa, J. M., Smith, L., Adams, Z. J. O., Khallaayoune, K., Sotiraki, S., Stefanakis, A., Farkas, R., & Ready, P. D. (2009). Molecular genetic analysis of populations of Wohlfahrt's wound myiasis fly, Wohlfahrtia magnifica, in outbreak populations from Greece and Morocco. *Medical and Veterinary Entomology*, *23*, 72–79. https://doi.org/10.1111/j.1365-2915.2009.00780.x

Hammond, A., Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., Gribble, M., Baker, D., Marois, E., Russell, S., Burt, A., Windbichler, N., Crisanti, A., & Nolan, T. (2016). A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nature Biotechnology*, *34*(1), 78–83. https://doi.org/10.1038/nbt.3439

He, P., Durand, N., & Dong, S. L. (2019). Insect olfactory proteins (from gene identification to functional characterization). *Frontiers in Physiology*, *10*, 1313. https://doi.org/10.3389/fphys.2019.01313

Hoff, K., Lomsadze, A., Borodovsky, M., & Stanke, M. (2019). Whole-genome annotation with BRAKER. *Methods in Molecular Biology*, *1962*, 65–95. https://doi.org/10.1007/978-1-4939-9173-0_5

Hu, X., Richtman, N. M., Zhao, J. Z., Duncan, K. E., Niu, X., Procyk, L. A., Oneal, M. A., Kernodle, B. M., Steimel, J. P., Crane, V. C., Sandahl, G., Ritland, J. L., Howard, R. J., Presnail, J. K., Lu, A. L., & Wu, G. (2016). Discovery of midgut genes for the RNA interference control of corn rootworm. *Scientific Reports*, *6*(1), 1–12. https://doi.org/10.1603/ice.2016.111203

i5K Consortium. (2013). The i5K initiative: Advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *Journal of Heredity*, *104*(5), 595–600. https://doi.org/10.1093/jhered/est050

İpek, D. N. S., Şaki, C. E., & Çay, M. (2012). The investigation of lipid peroxidation, anti-oxidant levels and some hematological parameters in sheep naturally infested with Wohlfahrtia magnifica larvae. *Veterinary Parasitology*, *187*(1–2), 112–118. https://doi.org/10.1016/j.vetpar.2011.12.018

Junqueira Alves, C., Yotoko, K., Zou, H., & Friedel, R. H. (2019). Origin and evolution of plexins, semaphorins, and met receptor tyrosine kinases. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-38512-y

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kim, Y. I., Nam, I. K., Um, J. Y., & Choe, S. K. (2019). Regulatory role of Wdr24 in autophagy activity during zebrafish embryogenesis. *Molecular & Cellular Toxicology*, *15*(1), 85–92. https://doi.org/10.1007/s13273-019-0010-3

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kingan, S. B., Heaton, H., Cudini, J., Lambert, C. C., Baybayan, P., Galvin, B. D., Durbin, R., Korlach, J., & MKN, L. (2019). A high-quality de novo genome assembly from a single mosquito using PacBio sequencing. *Genes*, *10*(1), 62. https://doi.org/10.3390/genes10010062

Knipling, E. F. (1955). Possibilities of insect control or eradication through the use of sexually sterile males. *Journal of Economic Entomology*, *48*(4), 459–462. https://doi.org/10.1093/jee/48.4.459

Kokcam, I., & Saki, C. E. (2005). A case of cutaneous myiasis caused by Wohlfahrtia magnifica. *The Journal of Dermatology*, *32*(6), 459–463. https://doi.org/10.1111/j.1346-8138.2005.tb00780.x

Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, *47*(D1), D807–D811. https://doi.org/10.1093/nar/gky1053

Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H. H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, *35*(9), 3100–3108. https://doi.org/10.1093/nar/gkm160

Li, H., Oyun, G., Bao, H., Yunzhang, L., Yang, B., Liu, T., & Demtu, E. (2020). Morphological and scanning electron microscopic (SEM) studies of the pupae of Wohlfahrtia magnifica. *Journal of Camel Practice and Research*, *27*(1), 17–22. https://doi.org/10.5958/2277-8934.2020.00003.x

Li, F., Wantuch, H. A., Linger, R. J., Belikoff, E. J., & Scott, M. J. (2014). Transgenic sexing system for genetic control of the Australian sheep blow fly Lucilia cuprina. *Insect Biochemistry and Molecular Biology*, *51*, 80–88. https://doi.org/10.1016/j.ibmb.2014.06.001

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., & Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, *33*(20), 6494–6506. https://doi.org/10.1093/nar/gki937

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, *25*(5), 955–964. https://doi.org/10.1093/nar/25.5.955

Martinez, R. I., Cruz, S. M. D., Rodriguez, R., Lopez, D. M., Parra, M. S., & Navio, F. A. (1987). Myiasis caused by Wohlfartia magnifica in southern Spain. *Israel Journal of Veterinary Medicine*, *43*(1), 34–41.

Martinson, E. O., Peyton, J., Kelkar, Y. D., Jennings, E. C., Benoit, J. B., Werren, J. H., & Denlinger, D. L. (2019). Genome and ontogenetic-based transcriptomic analyses of the flesh fly, Sarcophaga

bullata. *G3: Genes Genomes Genetics*, *9*(5), 1313–1320. https://doi.org/10.1534/g3.119.400148

Meng, F., Liu, Z., Han, H., Finkelbergs, D., Jiang, Y., Zhu, M., Wang, Y., Sun, Z., Chen, C., Guo, Y., & Cai, J. (2020). Chromosome-level genome assembly of Aldrichina grahami, a forensically important blowfly. *GigaScience*, *9*(3), giaa020. https://doi.org/10.1093/gigascience/giaa020

Moshaverinia, A., Moghaddas, E., Maleki, M., & Borji, H. (2013). Gingival myiasis of camel (Camelus dromedarius) caused by Wohlfahrtia magnifica. *Scientia Parasitologica*, *14*(2), 85–87.

Moussian, B. (2010). Recent advances in understanding mechanisms of insect cuticle differentiation. *Insect Biochemistry and Molecular Biology*, *40*(5), 363–375. https://doi.org/10.1016/j.ibmb.2010.03.003

Murmu, M. S., Stinnakre, J., & Martin, J. R. (2010). Presynaptic Ca2+ stores contribute to odor-induced responses in *Drosophila* olfactory receptor neurons. *Journal of Experimental Biology*, *213*(24), 4163–4173. https://doi.org/10.1242/jeb.046474

Olsen, L. C., & Faergeman, N. J. (2012). Chemical genomics and emerging DNA technologies in the identification of drug mechanisms and drug targets. *Current Topics in Medicinal Chemistry*, *12*(12), 1331–1345. https://doi.org/10.2174/156802612801319025

Overton, P. M., Meadows, L. A., Urban, J., & Russell, S. (2002). Evidence for differential and redundant function of the sox genes Dichaete and SoxN during CNS development in *Drosophila*. *Development*, *129*(18), 4219–4228. https://doi.org/10.1242/dev.129.18.4219

Papanicolaou, A., Schetelig, M. F., Arensburger, P., Atkinson, P. W., Benoit, J. B., Bourtzis, K., Castañera, P., Cavanaugh, J. P., Chao, H., Childers, C., Curril, I., Dinh, H., Doddapaneni, H., Dolan, A., Dugan, S., Friedrich, M., Gasperi, G., Geib, S., Georgakilas, G., … Handler, A. M. (2016). The whole genome sequence of the Mediterranean fruit fly, Ceratitis capitata (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biology*, *17*(1), 1–31. https://doi.org/10.1186/s13059-016-1049-2

Pertea, G., & Pertea, M. (2020). GFF utilities: GffRead and GffCompare. *F1000Research*, *9*, 304. https://doi.org/10.12688/f1000research.23297.2

Price, D. M., Jin, Z., Rabinovitch, S., & Campbell, S. D. (2002). Ectopic expression of the *Drosophila* Cdk1 inhibitory kinases, Wee1 and Myt1, interferes with the second mitotic wave and disrupts pattern formation during eye development. *Genetics*, *161*(2), 721–731. https://doi.org/10.1093/genetics/161.2.721

Remesar, S., Otero, J. L., Panadero, R., Díez-Baños, P., Díaz, P., García-Díos, D., Benito, A., Panadero, R., Morrondo, P., & López, C. (2022). Traumatic myiasis by Wohlfahrtia magnifica in sheep flocks from southeastern Spain: Prevalence and risk factors. *Medical and Veterinary Entomology*, *36*(1), 30–37. https://doi.org/10.1111/mve.12548

Ren, L., Shang, Y., Yang, L., Wang, S., Wang, X., Chen, S., Bao, Z., An, D., Meng, F., Cai, J., & Guo, Y. (2021). Chromosome-level de novo genome assembly of Sarcophaga peregrina provides insights into the evolutionary adaptation of flesh flies. *Molecular Ecology Resources*, *21*(1), 251–262. https://doi.org/10.1111/1755-0998.13246

Richards, S., & Murali, S. C. (2015). Best practices in insect genome sequencing: What works and what doesn't. *Current Opinion in Insect Science*, *7*, 1–7. https://doi.org/10.1016/j.cois.2015.02.013

Riga, M., Denecke, S., Livadaras, I., Geibel, S., Nauen, R., & Vontas, J. (2020). Development of efficient RNAi in Nezara viridula for use in insecticide target discovery. *Archives of Insect Biochemistry and Physiology*, *103*(3), e21650. https://doi.org/10.1002/arch.21650

Ruiz Martinez, I., Cruz, S. M. D., Perez, J. M., & Diaz, M. (1992). Larval development and mortality rate in the screwworm fly Wohlfahrtia magnifica (Schiner, 1862) (Diptera: Sarcophagidae). *Research and Reviews in Parasitology*, *52*, 27–32.

Ruiz Martinez, I., & Leclercq, M. (1994). Data on distribution of screwworm fly Wohlfahrtia magnifica (Schiner) in southwestern Europe (Diptera: Sarcophagidae). *Notes Fauniques de Gembloux*, *28*, 53–60.

Ruiz Martínez, I., Pérez Jiménez, J. M., & Cruz Mira, M. (1993). Epidemiology of wohlfahrtiosis in sheep and goats. *Investigación Agraria, Producción y Sanidad Animales*, *8*(3), 299–311.

Ruiz-Martínez, I., Soler-Cruz, M. D., Benítez-Rodríguez, R., Díaz-López, D. M., Muñoz-Parra, M. S., & Florido Navío, A. (1987). Myiasis caused by Wohlfahrtia magnifica (Schiner, 1862) (Diptera: Sarcophagidae) in southern Spain. *Irish Journal of Veterinary Medicine*, *43*, 34–41.

Sandeman, R. M., Levot, G. W., Heath, A. C. G., James, P. J., Greeff, J. C., Scott, M. J., Batterham, P., & Bowles, V. M. (2014). Control of the sheep blowfly in Australia and New Zealand–are we there yet? *International Journal for Parasitology*, *44*(12), 879–891. https://doi.org/10.1016/j.ijpara.2014.08.009

Schmidt-Nielsen, K., Schmidt-Nielsen, B., Jarnum, S. A., & Houpt, T. R. (1956). Body temperature of the camel and its relation to water economy. *American Journal of Physiology-Legacy Content*, *188*(1), 103–112. https://doi.org/10.1152/ajplegacy.1956.188.1.103

Schnorrer, F., Kalchhauser, I., & Dickson, B. J. (2007). The transmembrane protein Kon-tiki couples to Dgrip to mediate myotube targeting in drosophila. *Developmental Cell*, *12*(5), 751–766. https://doi.org/10.1016/j.devcel.2007.02.017

Schnur, H. J., Zivotofsky, D., & Wilamowski, A. (2009). Myiasis in domestic animals in Israel. *Veterinary Parasitology*, *161*(3–4), 352–355. https://doi.org/10.1016/j.vetpar.2009.01.026

Scholl, P. J., Colwell, D. D., & Cepeda-Palacios, R. (2019). Myiasis (Muscoidea, Oestroidea). In *Medical and veterinary entomology* (pp. 383–419). Academic Press. https://doi.org/10.1016/B978-0-12-814043-7.00019-4

Scott, M. J., Benoit, J. B., Davis, R. J., Bailey, S. T., Varga, V., Martinson, E. O., Hickner, P. V., Syed, Z., Cardoso, G. A., Torres, T. T., Weirauch, M. T., Scholl, E. H., Phillippy, A. M., Sagel, A., Vasquez, M., Quintero, G., & Skoda, S. R. (2020). Genomic analyses of a livestock pest, the New World screwworm, find potential targets for genetic control programs. *Communications Biology*, *3*(1), 1–14. https://doi.org/10.1038/s42003-020-01152-4

Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*, *11*(10), e0163962. https://doi.org/10.1371/journal.pone.0163962

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Soler Cruz, M. D., Vega Robles, M. C., Trapman, J. J., & Thomas, G. (1998). Comparative rearing of Wohlfahrtia magnifica (Diptera: Sarcophagidae) in dead and living tissues and the impact of cold storage on pupal survival. *Journal of Medical Entomology*, *35*, 153–156. https://doi.org/10.1093/jmedent/35.2.153

Soriano, N. S., & Russell, S. (1998). The drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline. *Development*, *125*(20), 3989–3996. https://doi.org/10.1242/dev.125.20.3989

Sotiraki, S., Farkas, R., & Hall, M. J. R. (2010). Fleshflies in the flesh: Epidemiology, population genetics and control of outbreaks of traumatic myiasis in the Mediterranean Basin. *Veterinary Parasitology*, *174*(1–2), 12–18. https://doi.org/10.1016/j.vetpar.2010.08.010

Sotiraki, S., Martin, J. R., & Hall, M. R. J. (2012). A review of comparative aspects of myiasis in goats and sheep in Europe. *Small Ruminant Research*, *103*, 75–83. https://doi.org/10.1016/j.smallrumres.2011.10.021

Sotiraki, S., Stefanakis, A., Hall, M. J. R., Farkas, R., & Graf, J. F. (2005). Wohlfahrtiosis in sheep and the role of dicyclanil in its prevention. *Veterinary Parasitology*, *131*(1–2), 107–117. https://doi.org/10.1016/j.vetpar.2005.04.026

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stanke, M., Steinkamp, R., Waack, S., & Morgenstern, B. (2004). AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Research*, *32*(suppl_2), W309–W312.

Szpila, K., Hall, M. J. R., Wardhana, A. H., & Pape, T. (2014). Morphology of the first instar larva of obligatory traumatic myiasis agents (Diptera: Calliphoridae, Sarcophagidae). *Parasitology Research*, *113*(5), 1629–1640. https://doi.org/10.1007/s00436-014-3808-x

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. https://doi.org/10.1080/10635150701472164

Valentin, A., Baumann, M. P. O., Schein, E., & Bajanbileg, S. (1997). Genital myiasis (Wohlfahrtiosis) in camel herds of Mongolia. *Veterinary Parasitology*, *73*(3–4), 335–346. https://doi.org/10.1016/s0304-4017(97)00127-1

Wyss, J. H. (2006). Screwworm eradication in the Americas. *Annals of the New York Academy of Sciences*, *916*(1), 186–193. https://doi.org/10.1111/j.1749-6632.2000.tb05289.x

Yan, L., Zhang, M., Tang, L., Ente, M., Ma, X., Chu, H., Li, K., Hu, D., & Zhang, D. (2019). First reports of nasal and traumatic myiasis infection in endangered Przewalski's horses (Equus ferus przewalskii). *International Journal for Parasitology: Parasites and Wildlife*, *9*, 21–24. https://doi.org/10.1016/j.ijppaw.2019.03.018

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Yasuda, M. (1940). Morphology of the larva of Wohlfahrtia magnifica Schin. Found in a wound on a camel in Inner Mongolia. *Journal of the Chosen Natural History Society*, *7*(29), 27–36.

Ye, Y. X., Zhang, H. H., Li, D. T., Zhuo, J. C., Shen, Y., Hu, Q. L., & Zhang, C. X. (2021). Chromosome-level assembly of the brown planthopper genome with a characterized Y chromosome. *Molecular Ecology Resources*, *21*(4), 1287–1298. https://doi.org/10.1111/1755-0998.13328

Zhang, Z., Xiao, J., Wu, J., Zhang, H., Liu, G., Wang, X., & Dai, L. (2012). ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochemical and Biophysical Research Communications*, *419*(4), 779–781. https://doi.org/10.1016/j.bbrc.2012.02.101

Zumpt, F. (1965). *Myiasis in man and animals in the Old World: A textbook for physicians, veterinarians and zoologists*. Butterworths.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Jia, Z., Hasi, S., Vogl, C., & Burger, P. A. (2022). Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates. *Molecular Ecology Resources*, *22*, 2744–2757. https://doi.org/10.1111/1755-0998.13654

Supplementary Materials for:

Genomic insights into evolution and control of *Wohlfahrtia magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates

Zhipeng Jia[1], Surong Hasi[2], Claus Vogl[3], Pamela A. Burger[1*]

1. Department of Interdisciplinary Life Sciences, Research Institute of Wildlife Ecology, University of Veterinary Medicine Vienna, Vienna, Austria
2. Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Hohhot, China
3. Department of Biomedical Sciences, Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna, Vienna, Austria
*Corresponding author: pamela.burger@vetmeduni.ac.at

The following contents are included:

Supplementary Tables S1 to S6
Supplementary Data S1

Table S1. Detection results of DNA sample for low input library preparation.

Sample	Fluorescence concentration (ng/μl)	UV concentration (ng/μl)	Volume(μl)	OD260/280	OD260/230	Total amount (μg)
A single adult *W. magnifica*	27	30.10	45	1.77	2.11	1.215

Table S2. BUSCO analyses of *W. magnifica* and other nine dipterous flies based on diptera_odb10 gene set (number of BUSCOs: 3285). Genome: assessment of genome assembly completeness; gene set: assessment of predicted gene set completeness.

Species	Mode	Complete (Single Copy)		Complete (Duplicated)		Fragmented		Missing	
		Percentage	Numbers	Percentage	Numbers	Percentage	Numbers	Percentage	Numbers
W. magnifica	Genome	98.2	3,226	0.6	19	0.6	19	0.6	21
	Gene set	81.9	2,691	15.3	501	1.3	42	1.5	51
A.aegypti	Genome	94.4	3,101	3.1	101	1.1	36	1.4	47
	Gene set	60.5	1,989	38.9	1,278	0.2	7	0.5	11
A. gambiae	Genome	97.5	3,204	1.7	56	0.1	4	0.7	21
	Gene set	87.5	2,874	11.6	381	0.4	13	0.5	17
D. melanogaster	Genome	99.0	3,252	0.4	14	0.2	5	0.4	14
	Gene set	41.8	1,372	58.2	1,913	0.0	0	0.0	0
G. morsitans	Genome	96.7	3,176	1.3	43	1.0	32	1.0	34
	Gene set	89.8	2,949	2.3	75	1.9	61	6.0	200
L. cuprina	Genome	96.2	3,159	1.6	53	0.9	31	1.3	42
	Gene set	87.5	2,876	1.7	56	1.2	41	9.6	312
M. destructor	Genome	85.3	2,803	1.4	45	4.3	142	9.0	295
	Gene set	71.1	2,334	13.6	448	4.4	146	10.9	357
M. domestica	Genome	94.9	3,116	2.2	71	1.2	39	1.7	59
	Gene set	73.2	2,406	25.7	845	0.8	25	0.3	9
S. bullata	Genome	86.7	2,847	0.8	27	6.8	225	5.7	186
	Gene set	77.9	2560	1.1	35	8.7	286	12.3	404
S. calcitrans	Genome	96.2	3,159	0.9	29	1.1	37	1.8	60
	Gene set	67.7	2,223	31.6	1,039	0.5	15	0.2	8

Table S3. Comparison of summary statistics of repeat element of *W. magnifica* and other dipterous flies.

Repeat element	*W. magnifica*			*C. hominivorax*			*L. cuprina*			*S. bullata*		
	Numbers	Bases	% of Genome	Numbers	Bases	% of Genome	Numbers	Bases	% of Genome	Numbers	Bases	% of Genome
SINEs	0	0	0.00	63	5,637	0.00	434	64,889	0.01	33,478	7,221,604	1.52
LINEs	313,495	97,444,031	12.92	/	12,057,930	2.26	61,619	15,766,996	3.35	179,696	32,921,812	6.91
LTR	98,752	29,004,031	3.85	/	2,848,352	0.53	16,688	5,901,984	1.25	17,235	6,433,260	1.35
DNA	473,914	125,334,368	16.62	/	36,993,232	6.93	60,359	12,837,946	2.73	127,299	19,969,616	4.19
Unclassified	976,699	176,211,233	23.37	669,162	135,021,258	25.28	383,500	78,482,695	16.68	397,153	59,304,762	12.45
Simple repeats	279,951	18,690,856	2.48	320,870	17,432,637	3.26	150,125	24,344,465	5.17	416,660	18,229,578	3.83
Low complexity	61,088	2,981,425	0.40	73,844	3,798,472	0.71	370,031	18,958,378	4.03	89,350	4,553,468	0.96
Total	/	450,243,625	59.71	1,467,371	241,520,001	45.22	1,058,216	272,180,260	57.82	/	148,389,050	31.15

Table S4. Distribution of genes in *W. magnifica* and other dipterous flies. 1:1 orthologs: single-copy orthologs; N:N orthologs: multiple-copy orthologs; species-specific orthologs: present in specific species including multi copy and single copy; other orthologs: the remaining orthologs.

Type		*A. aegypti*	*A. gambiae*	*D. melanogaster*	*G. morsitans*	*L. cuprina*	*M. destructor*	*M. domestica*	*S. bullata*	*S. calcitrans*	*W. magnifica*
Number of genes		14,718	13,094	13,968	12,494	14,452	19,927	14,402	15,763	14,078	16,718
1:1 orthologs		2,045	2,045	2,045	2,045	2,045	2,045	2,045	2,045	2,045	2,045
N:N orthologs		5,240	4,442	4,604	4,725	4,352	4,544	5,137	5,199	4,895	4,972
Species-specific orthologs	Multi copy	1,029	834	707	402	723	4,259	317	501	289	943
	Single copy	456	563	1261	1,099	2,048	5,066	319	2,459	320	793
Other orthologs		5,948	5,210	5,351	4,223	5,284	4,013	6,584	5,559	6,529	7,965

Table S5. GO classification of expanded gene families of *W. magnifica*.

GO_ID	GO_class	GO_description	Pvalue	Qvalue	Gene Numbers	Reference Numbers	RichFactor
GO:0005214	Molecular Function	structural constituent of chitin-based cuticle	1.14E-54	1.17E-52	73	140	0.521
GO:0042302	Molecular Function	structural constituent of cuticle	1.37E-52	6.96E-51	74	152	0.487
GO:0005576	Cellular Component	extracellular region	9.71E-42	1.10E-39	95	453	0.21
GO:0044421	Cellular Component	extracellular region part	3.79E-40	2.14E-38	88	400	0.22
GO:0042335	Biological Process	cuticle development	1.61E-38	1.10E-35	73	202	0.361
GO:0008023	Cellular Component	transcription elongation factor complex	1.71E-35	6.44E-34	38	65	0.585
GO:0005198	Molecular Function	structural molecule activity	3.97E-33	1.35E-31	84	337	0.249
GO:0001824	Biological Process	blastocyst development	8.40E-34	2.89E-31	27	27	1
GO:0000785	Cellular Component	chromatin	1.03E-32	2.91E-31	56	186	0.301
GO:0032784	Biological Process	regulation of DNA-templated transcription, elongation	3.98E-33	9.12E-31	38	57	0.667
GO:0031062	Biological Process	positive regulation of histone methylation	2.13E-29	3.66E-27	31	42	0.738
GO:0018023	Biological Process	peptidyl-lysine trimethylation	1.25E-26	1.72E-24	27	35	0.771
GO:0044427	Cellular Component	chromosomal part	1.25E-24	2.81E-23	59	288	0.205
GO:0032786	Biological Process	positive regulation of DNA-templated transcription, elongation	1.15E-24	1.32E-22	28	42	0.667
GO:0001701	Biological Process	in utero embryonic development	1.65E-24	1.62E-22	27	39	0.692
GO:0031060	Biological Process	regulation of histone methylation	2.57E-23	2.21E-21	31	57	0.544
GO:0060795	Biological Process	cell fate commitment involved in formation of primary germ layer	2.17E-22	1.66E-20	27	44	0.614
GO:0006354	Biological Process	DNA-templated transcription, elongation	3.50E-22	2.41E-20	38	95	0.4
GO:0031058	Biological Process	positive regulation of histone modification	3.89E-22	2.43E-20	31	61	0.508
GO:1903310	Biological Process	positive regulation of chromatin modification	7.35E-22	4.21E-20	31	62	0.5
GO:2001252	Biological Process	positive regulation of chromosome organization	2.50E-21	1.32E-19	31	64	0.484
GO:0017085	Biological Process	response to insecticide	4.23E-21	2.08E-19	25	40	0.625
GO:0001704	Biological Process	formation of primary germ layer	3.71E-18	1.70E-16	27	58	0.466
GO:0018022	Biological Process	peptidyl-lysine methylation	4.10E-18	1.76E-16	28	63	0.444
GO:0043414	Biological Process	macromolecule methylation	5.43E-17	2.20E-15	38	128	0.297

GO:0016571	Biological Process	histone methylation	1.98E-16	7.16E-15	31	88	0.352
GO:0031056	Biological Process	regulation of histone modification	1.98E-16	7.16E-15	31	88	0.352
GO:0005694	Cellular Component	chromosome	3.97E-16	7.47E-15	59	421	0.14
GO:0032259	Biological Process	methylation	3.92E-16	1.35E-14	38	135	0.281
GO:0006325	Biological Process	chromatin organization	6.40E-16	2.10E-14	67	379	0.177
GO:0043566	Molecular Function	structure-specific DNA binding	9.93E-16	2.53E-14	31	98	0.316
GO:1902275	Biological Process	regulation of chromatin organization	1.62E-15	5.06E-14	34	113	0.301
GO:0044451	Cellular Component	nucleoplasm part	3.32E-15	5.36E-14	49	316	0.155
GO:0005654	Cellular Component	nucleoplasm	3.83E-15	5.41E-14	50	329	0.152
GO:0010628	Biological Process	positive regulation of gene expression	1.97E-15	5.90E-14	41	163	0.252
GO:0006479	Biological Process	protein methylation	2.35E-15	6.74E-14	31	95	0.326
GO:0014031	Biological Process	mesenchymal cell development	4.87E-15	1.34E-13	23	52	0.442
GO:0048762	Biological Process	mesenchymal cell differentiation	8.12E-15	2.15E-13	23	53	0.434
GO:0009636	Biological Process	response to toxic substance	8.57E-15	2.18E-13	26	69	0.377
GO:0008213	Biological Process	protein alkylation	2.18E-14	5.37E-13	31	102	0.304
GO:0033044	Biological Process	regulation of chromosome organization	2.38E-14	5.65E-13	37	144	0.257
GO:1903308	Biological Process	regulation of chromatin modification	2.95E-14	6.76E-13	31	103	0.301
GO:0008593	Biological Process	regulation of Notch signaling pathway	3.97E-14	8.81E-13	31	104	0.298
GO:0007369	Biological Process	gastrulation	7.17E-13	1.54E-11	28	94	0.298
GO:0048864	Biological Process	stem cell development	1.06E-12	2.21E-11	24	70	0.343
GO:0045893	Biological Process	positive regulation of transcription, DNA-templated	1.35E-12	2.59E-11	29	103	0.282
GO:1902680	Biological Process	positive regulation of RNA biosynthetic process	1.35E-12	2.59E-11	29	103	0.282
GO:1903508	Biological Process	positive regulation of nucleic acid-templated transcription	1.35E-12	2.59E-11	29	103	0.282
GO:0060485	Biological Process	mesenchyme development	1.48E-12	2.75E-11	23	65	0.354
GO:0051254	Biological Process	positive regulation of RNA metabolic process	2.18E-12	3.96E-11	31	119	0.261
GO:0007219	Biological Process	Notch signaling pathway	5.64E-12	9.95E-11	31	123	0.252
GO:0009617	Biological Process	response to bacterium	6.69E-12	1.15E-10	52	305	0.17
GO:0043009	Biological Process	chordate embryonic development	4.77E-11	8.01E-10	28	110	0.255
GO:0006259	Biological Process	DNA metabolic process	7.78E-11	1.27E-09	65	460	0.141

GO:0042742	Biological Process	defense response to bacterium	1.19E-10	1.91E-09	39	205	0.19
GO:0010557	Biological Process	positive regulation of macromolecule biosynthetic process	1.60E-10	2.50E-09	29	123	0.236
GO:0010638	Biological Process	positive regulation of organelle organization	1.67E-10	2.55E-09	31	139	0.223
GO:0045935	Biological Process	positive regulation of nucleobase-containing compound metabolic process	1.97E-10	2.95E-09	32	148	0.216
GO:0051173	Biological Process	positive regulation of nitrogen compound metabolic process	4.87E-10	7.13E-09	32	153	0.209
GO:0031328	Biological Process	positive regulation of cellular biosynthetic process	6.56E-10	9.40E-09	29	130	0.223
GO:0014074	Biological Process	response to purine-containing compound	1.43E-09	2.01E-08	14	32	0.438
GO:0098542	Biological Process	defense response to other organism	1.58E-09	2.17E-08	39	223	0.175
GO:0051276	Biological Process	chromosome organization	4.88E-09	6.58E-08	67	530	0.126
GO:0002385	Biological Process	mucosal immune response	6.35E-09	8.40E-08	9	13	0.692
GO:0009891	Biological Process	positive regulation of biosynthetic process	1.55E-08	2.02E-07	29	148	0.196
GO:0018205	Biological Process	peptidyl-lysine modification	1.73E-08	2.20E-07	28	140	0.2
GO:0000786	Cellular Component	nucleosome	3.37E-08	3.81E-07	7	10	0.7
GO:0044815	Cellular Component	DNA packaging complex	3.37E-08	3.81E-07	7	10	0.7
GO:0048863	Biological Process	stem cell differentiation	3.19E-08	3.99E-07	24	110	0.218
GO:0031012	Cellular Component	extracellular matrix	5.05E-08	5.18E-07	17	80	0.212
GO:0090304	Biological Process	nucleic acid metabolic process	4.48E-08	5.50E-07	166	1884	0.088
GO:0051707	Biological Process	response to other organism	7.46E-08	9.01E-07	53	405	0.131
GO:0043207	Biological Process	response to external biotic stimulus	8.80E-08	1.04E-06	53	407	0.13
GO:0009607	Biological Process	response to biotic stimulus	1.22E-07	1.42E-06	53	411	0.129
GO:0002168	Biological Process	instar larval development	1.27E-07	1.46E-06	10	21	0.476
GO:0009952	Biological Process	anterior/posterior pattern specification	2.02E-07	2.27E-06	23	112	0.205
GO:0000122	Biological Process	negative regulation of transcription from RNA polymerase II promoter	2.20E-07	2.44E-06	10	22	0.455
GO:0002251	Biological Process	organ or tissue specific immune response	3.27E-07	3.58E-06	9	18	0.5
GO:0031401	Biological Process	positive regulation of protein modification process	4.87E-07	5.23E-06	36	242	0.149
GO:0016568	Biological Process	chromatin modification	1.12E-06	1.18E-05	41	304	0.135
GO:0034605	Biological Process	cellular response to heat	1.27E-06	1.33E-05	6	8	0.75
GO:0043228	Cellular Component	non-membrane-bounded organelle	3.92E-06	3.37E-05	100	1507	0.066

GO:0043232	Cellular Component	intracellular non-membrane-bounded organelle	3.92E-06	3.37E-05	100	1507	0.066
GO:0032993	Cellular Component	protein-DNA complex	4.17E-06	3.37E-05	7	17	0.412
GO:0045892	Biological Process	negative regulation of transcription, DNA-templated	3.38E-06	3.42E-05	17	78	0.218
GO:1903507	Biological Process	negative regulation of nucleic acid-templated transcription	3.38E-06	3.42E-05	17	78	0.218
GO:0051130	Biological Process	positive regulation of cellular component organization	3.65E-06	3.64E-05	31	210	0.148
GO:0015935	Cellular Component	small ribosomal subunit	6.56E-06	4.94E-05	10	40	0.25
GO:0018193	Biological Process	peptidyl-amino acid modification	1.01E-05	9.95E-05	28	189	0.148
GO:0003676	Molecular Function	nucleic acid binding	4.90E-06	0.000100047	120	1449	0.083
GO:0010604	Biological Process	positive regulation of macromolecule metabolic process	1.36E-05	0.000131563	49	430	0.114
GO:0032270	Biological Process	positive regulation of cellular protein metabolic process	1.41E-05	0.000135208	36	279	0.129
GO:0051247	Biological Process	positive regulation of protein metabolic process	1.66E-05	0.000156402	36	281	0.128
GO:0009792	Biological Process	embryo development ending in birth or egg hatching	1.80E-05	0.0001674	30	216	0.139
GO:0006260	Biological Process	DNA replication	2.82E-05	0.000258821	19	109	0.174
GO:0016569	Biological Process	covalent chromatin modification	3.09E-05	0.000276442	31	233	0.133
GO:0016570	Biological Process	histone modification	3.09E-05	0.000276442	31	233	0.133
GO:0006139	Biological Process	nucleobase-containing compound metabolic process	3.15E-05	0.000278039	167	2094	0.08
GO:1902679	Biological Process	negative regulation of RNA biosynthetic process	4.56E-05	0.000397085	17	94	0.181
GO:0035821	Biological Process	modification of morphology or physiology of other organism	5.65E-05	0.000486037	8	24	0.333
GO:0034728	Biological Process	nucleosome organization	6.18E-05	0.000525241	13	61	0.213
GO:0042594	Biological Process	response to starvation	8.42E-05	0.000706588	10	39	0.256
GO:0046483	Biological Process	heterocycle metabolic process	0.000102238	0.000847468	167	2137	0.078
GO:0006952	Biological Process	defense response	0.000110019	0.000901105	45	416	0.108
GO:2000113	Biological Process	negative regulation of cellular macromolecule biosynthetic process	0.000149171	0.001207409	17	103	0.165
GO:0006306	Biological Process	DNA methylation	0.000167211	0.00132231	7	21	0.333
GO:0044728	Biological Process	DNA methylation or demethylation	0.000167211	0.00132231	7	21	0.333

GO:0006323	Biological Process	DNA packaging	0.000203612	0.001591876	10	43	0.233
GO:0006305	Biological Process	DNA alkylation	0.00031675	0.002421379	7	23	0.304
GO:0006334	Biological Process	nucleosome assembly	0.00031675	0.002421379	7	23	0.304
GO:0031981	Cellular Component	nuclear lumen	0.000344006	0.002429541	64	958	0.067
GO:0030246	Molecular Function	carbohydrate binding	0.000156879	0.002666943	7	22	0.318
GO:0006725	Biological Process	cellular aromatic compound metabolic process	0.00038624	0.002920147	167	2190	0.076
GO:0048029	Molecular Function	monosaccharide binding	0.000210907	0.003073213	5	11	0.455
GO:0043189	Cellular Component	H4/H2A histone acetyltransferase complex	0.000495155	0.003108475	6	24	0.25
GO:1902562	Cellular Component	H4 histone acetyltransferase complex	0.000495155	0.003108475	6	24	0.25
GO:0031514	Cellular Component	motile cilium	0.000644736	0.003834481	4	10	0.4
GO:0070013	Cellular Component	intracellular organelle lumen	0.000700343	0.003904975	64	983	0.065
GO:0044449	Cellular Component	contractile fiber part	0.000737676	0.003904975	11	80	0.138
GO:0043233	Cellular Component	organelle lumen	0.000760261	0.003904975	64	986	0.065
GO:0006304	Biological Process	DNA modification	0.000558147	0.00417397	7	25	0.28
GO:0051253	Biological Process	negative regulation of RNA metabolic process	0.000573229	0.004240661	17	115	0.148
GO:0033043	Biological Process	regulation of organelle organization	0.000703057	0.005145779	37	350	0.106
GO:0031974	Cellular Component	membrane-enclosed lumen	0.001291911	0.006347214	64	1006	0.064
GO:0043292	Cellular Component	contractile fiber	0.001360206	0.006404302	11	86	0.128
GO:1901360	Biological Process	organic cyclic compound metabolic process	0.000899214	0.006493923	170	2273	0.075
GO:0071824	Biological Process	protein-DNA complex subunit organization	0.000906129	0.006493923	13	79	0.165
GO:0016779	Molecular Function	nucleotidyltransferase activity	0.00055632	0.007093077	12	70	0.171
GO:0034641	Biological Process	cellular nitrogen compound metabolic process	0.001156246	0.008201002	173	2331	0.074
GO:0009408	Biological Process	response to heat	0.001214106	0.008523522	8	36	0.222
GO:0003677	Molecular Function	DNA binding	0.000902933	0.00994628	49	544	0.09
GO:0016628	Molecular Function	oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	0.000975125	0.00994628	8	37	0.216
GO:0010558	Biological Process	negative regulation of macromolecule biosynthetic process	0.001491014	0.010361791	17	125	0.136
GO:0006357	Biological Process	regulation of transcription from RNA polymerase II promoter	0.001621179	0.01098685	52	564	0.092
GO:0051262	Biological Process	protein tetramerization	0.001625076	0.01098685	6	22	0.273

GO:0045934	Biological Process	negative regulation of nucleobase-containing compound metabolic process	0.001628864	0.01098685	17	126	0.135
GO:0031497	Biological Process	chromatin assembly	0.001811402	0.012099465	7	30	0.233
GO:0031327	Biological Process	negative regulation of cellular biosynthetic process	0.002108493	0.01394849	17	129	0.132
GO:0006333	Biological Process	chromatin assembly or disassembly	0.002693413	0.01764827	7	32	0.219
GO:0006366	Biological Process	transcription from RNA polymerase II promoter	0.003061846	0.019873114	52	581	0.09
GO:0009890	Biological Process	negative regulation of biosynthetic process	0.003701546	0.023751898	17	136	0.125
GO:0072319	Biological Process	vesicle uncoating	0.003728496	0.023751898	2	2	1
GO:0050919	Biological Process	negative chemotaxis	0.004071389	0.025698307	6	26	0.231
GO:0009893	Biological Process	positive regulation of metabolic process	0.004396204	0.027496257	49	550	0.089
GO:0009266	Biological Process	response to temperature stimulus	0.004902755	0.030388247	15	117	0.128
GO:0006406	Biological Process	mRNA export from nucleus	0.006016631	0.03663223	5	20	0.25
GO:0071427	Biological Process	mRNA-containing ribonucleoprotein complex export from nucleus	0.006016631	0.03663223	5	20	0.25
GO:0006342	Biological Process	chromatin silencing	0.007356202	0.043629889	7	38	0.184
GO:0045814	Biological Process	negative regulation of gene expression, epigenetic	0.007356202	0.043629889	7	38	0.184
GO:0065004	Biological Process	protein-DNA complex assembly	0.007356202	0.043629889	7	38	0.184
GO:0043933	Biological Process	macromolecular complex subunit organization	0.00757468	0.044541705	80	1010	0.079

Table S6. Genes with evidence of positive selection in the genome of *W. magnifica*.

Gene.ID	*D. Melanogaster* orthologs	Annotated function of *D. melanogaster* orthologs	l0	l1	2(l1-l0)	pvalue	p.adj
Woma_00015236-RA	FBgn0025741	Gene=Plex A, Plexin A	36626.655076	36696.497668	139.685	0.000e+00	0.000000e+00
Woma_00006011-RB	FBgn0051158	Gene=Efa6, Exchange factor for Arf 6	18461.556751	18487.733199	52.3528	4.637e-13	1.174707e-10
Woma_00009437-RB	FBgn0041789	Gene=Pax, Paxillin	-7267.258270	-7289.974646	45.4328	1.580e-11	3.002000e-09
Woma_00012634-RA	FBgn0004360	Gene=Wnt2, Wnt oncogene analog 2	-6392.824908	-6410.287286	34.9248	3.427e-09	5.209040e-07
Woma_00005680-RA	FBgn0037648	Gene=CG11975	-6113.242801	-6128.605055	30.7246	2.974e-08	3.767067e-06
Woma_00006399-RA	FBgn0035532	Gene=CG15014	-6957.565389	-6972.300373	29.47	5.679e-08	6.165771e-06
Woma_00007179-RA	FBgn0011286	Gene=RyR, Ryanodine receptor	79959.098831	79972.587268	26.9768	2.059e-07	1.956050e-05
Woma_00010254-RA	FBgn0020251	Gene=sfl, sulfateless	13872.742861	13884.719202	23.9526	9.874e-07	8.338044e-05
Woma_00004948-RA	FBgn0032901	Gene=sky, skywalker	-6132.164444	-6142.896530	21.4642	3.605e-06	2.689709e-04
Woma_00000964-RA	FBgn0033482	Gene=CG1371	28464.907561	28475.566048	21.317	3.893e-06	2.689709e-04
Woma_00012285-RA	FBgn0033166	Gene=Eaf, ELL-associated factor	-1396.137382	-1406.312657	20.3506	6.447e-06	4.083100e-04
Woma_00006976-RA	FBgn0000411	Gene=D, Dichaete	-4800.598853	-4810.275906	19.3541	1.086e-05	6.348923e-04
Woma_00012332-RA	FBgn0053111	Gene=CG33111	-8013.737171	-8022.924721	18.3751	1.814e-05	9.847429e-04
Woma_00010202-RA	FBgn0040298	Gene=Myt1, Myt1	10372.732751	10381.357864	17.2502	3.277e-05	1.660347e-03
Woma_00010517-RA	FBgn0036462	Gene=mRpL39, mitochondrial ribosomal protein L39	-7331.936570	-7339.708681	15.5442	8.060e-05	3.828500e-03
Woma_00009906-RA	FBgn0030246	Gene=CG1582	24202.940356	24210.037570	14.1944	1.649e-04	6.652000e-03
Woma_00009502-RA	FBgn0031878	Gene=sip2, septin interacting protein 2	11211.529651	11218.618818	14.1783	1.663e-04	6.652000e-03
Woma_00004111-RA	FBgn0031628	Gene=CG3294	11451.435441	11458.361536	13.8522	1.978e-04	7.516400e-03
Woma_00009853-RA	FBgn0038272	Gene=Dph2, Diphthamide biosynthesis 2	11121.555612	11128.296688	13.4822	2.408e-04	8.714667e-03
Woma_00001951-RA	FBgn0259223	Gene=CG42323	-2276.150447	-2282.501176	12.7015	3.654e-04	1.214017e-02
Woma_00010582-RA	FBgn0264494	Gene=CG17646	-2363.568276	-2369.913759	12.691	3.674e-04	1.214017e-02
Woma_00009603-RA	FBgn0028484	Gene=Ack, Activated Cdc42 kinase	13925.010681	13931.286369	12.5514	3.959e-04	1.253683e-02
Woma_00003792-RA	FBgn0037632	Gene=CCT7, Chaperonin containing TCP1 subunit 7	-4774.823590	-4781.005339	12.3635	4.378e-04	1.330912e-02
Woma_00005239-RA	FBgn0260940	Gene=lsn, larsen	-4603.138134	-4609.122887	11.9695	5.408e-04	1.580800e-02
Woma_00010013-RA	FBgn0034816	Gene=CG3085	-9131.729534	-9137.643070	11.8271	5.837e-04	1.643007e-02
Woma_00005272-RA	FBgn0043455	Gene=CG5986	-4759.062292	-4764.825344	11.5261	6.863e-04	1.798579e-02
Woma_00008799-RB	FBgn0284408	Gene=trol, terribly reduced optic lobes	64876.210451	64881.897841	11.3748	7.445e-04	1.886067e-02
Woma_00005324-RB	FBgn0260468	Gene=CG7950	-2446.441525	-2452.015719	11.1484	8.410e-04	2.061806e-02
Woma_00004396-RA	FBgn0027518	Gene=Wdr24, WD repeat domain 24	15946.376831	15951.694069	10.6345	1.110e-03	2.556364e-02
Woma_00015317-RA	FBgn0259994	Gene=OtopLa, Otopetrin-like a	-5843.961134	-5849.198485	10.4747	1.210e-03	2.671122e-02
Woma_00007237-RB	FBgn0038830	Gene=CG17272	-2484.239895	-2489.435794	10.3918	1.266e-03	2.671122e-02

Woma_00005723-RA	FBgn0032683	Gene=kon, Kon-tiki	49093.394907	49098.542078	10.2943	1.334e-03	2.671122e-02
Woma_00011309-RA	FBgn0027889	Gene=ball, ballchen	10865.829363	10870.950115	10.2415	1.373e-03	2.671122e-02
Woma_00006898-RA	FBgn0036504	Gene=yellow-k, yellow-k	-7672.990773	-7678.091869	10.2022	1.403e-03	2.671122e-02
Woma_00005570-RA	FBgn0039688	Gene=Kul, Kuzbanian-like	14059.741661	14064.838534	10.1937	1.409e-03	2.671122e-02
Woma_00000912-RA	FBgn0086359	Gene=Invadolysin, Invadolysin	10880.476650	10885.566905	10.1805	1.419e-03	2.671122e-02
Woma_00007702-RA	FBgn0003416	Gene=sl, small wing	11335.580908	11340.657219	10.1526	1.441e-03	2.671122e-02
Woma_00012533-RB	FBgn0033061	SmydA-5, SET and MYND domain containing, arthropod-specific, member 5	13332.982019	13337.935330	9.90662	1.647e-03	2.905273e-02
Woma_00000675-RA	FBgn0037218	Gene=aux, Auxilin	-8169.459507	-8174.393309	9.8676	1.682e-03	2.905273e-02
Woma_00003748-RA	FBgn0037881	Gene=GCC88, GRIP and coiled-coil domain containing 88 kDa	15179.276711	15184.104795	9.65616	1.887e-03	3.026041e-02
Woma_00012353-RB	FBgn0039528	Gene=dsd, distracted	24477.794386	24482.606668	9.62456	1.920e-03	3.026041e-02
Woma_00010519-RA	FBgn0052190	Gene=NUCB1, NUCB1	-8681.032986	-8685.830708	9.59544	1.951e-03	3.026041e-02
Woma_00005480-RB	FBgn0038065	Gene=Snx3, Sorting nexin 3	-2477.376080	-2481.989967	9.22778	2.384e-03	3.548615e-02
Woma_00007943-RA	FBgn0037376	Gene=Hat1, Histone acetyltransferase 1	-9415.611997	-9420.223187	9.22238	2.391e-03	3.548615e-02
Woma_00012568-RA	FBgn0050010	Gene=CG30010	-4536.353786	-4540.627293	8.54702	3.461e-03	4.962943e-02

Data S1. Parasitism-related genes.

ID of parasitism-related genes	Subject annotation
Woma_00015663-RA	XP_023298301.1 claw keratin-like [Lucilia cuprina]
Woma_00014139-RA	CAP09075.1 minos transposase [Drosophila hydei]
Woma_00014126-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00014123-RA	AAB04627.1 reverse transcriptase [Chironomus tentans]
Woma_00014158-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00014219-RA	XP_013104093.1 PREDICTED: fibrous sheath CABYR-binding protein-like [Stomoxys calcitrans]
Woma_00014325-RA	XP_023296566.1 zinc finger FYVE domain-containing protein 1-like isoform X1 [Lucilia cuprina]
Woma_00014349-RA	XP_023297127.1 flocculation protein FLO11-like [Lucilia cuprina]
Woma_00014366-RA	CAB39733.1 protease, reverse transcriptase, ribonuclease H, integrase, partial [Drosophila buzzatii]
Woma_00014504-RA	XP_023302563.1 chymotrypsin-2-like, partial [Lucilia cuprina]
Woma_00015569-RB	XP_011213444.1 PREDICTED: glucose-induced degradation protein 4 homolog [Bactrocera dorsalis]
Woma_00014567-RA	KNC28259.1 Coiled-coil domain-containing protein lobo [Lucilia cuprina]
Woma_00014578-RB	XP_023034128.1 endoplasmic reticulum metallopeptidase 1 [Drosophila willistoni]
Woma_00014628-RA	KNC29657.1 putative RNA-directed DNA polymerase from transposon X-element [Lucilia cuprina]
Woma_00014722-RA	AID61232.1 odorant receptor [Calliphora stygia]
Woma_00014743-RA	XP_023308343.1 RING-H2 finger protein ATL48-like [Lucilia cuprina]
Woma_00014772-RA	PIK48899.1 Retrovirus-related Pol polyprotein from transposon [Apostichopus japonicus]
Woma_00014785-RA	KNC33626.1 putative gustatory receptor 59d [Lucilia cuprina]
Woma_00014786-RA	XP_023307903.1 putative gustatory receptor 59d [Lucilia cuprina]
Woma_00014787-RA	XP_023307903.1 putative gustatory receptor 59d [Lucilia cuprina]
Woma_00008420-RB	XP_023299645.1 thrombospondin type-1 domain-containing protein 4-like [Lucilia cuprina]
Woma_00008466-RA	XP_013104572.1 PREDICTED: angiotensin-converting enzyme [Stomoxys calcitrans]
Woma_00008468-RA	XP_005187540.1 PREDICTED: SH3 domain-containing protein C23A1.17 [Musca domestica]
Woma_00008476-RA	XP_013102621.1 PREDICTED: centaurin-gamma-1A isoform X1 [Stomoxys calcitrans]
Woma_00008520-RA	XP_023294436.1 cytochrome c oxidase subunit 5B, mitochondrial-like [Lucilia cuprina]
Woma_00008507-RA	XP_023301793.1 zinc finger protein weckle-like [Lucilia cuprina]
Woma_00008485-RA	XP_005192200.2 PREDICTED: centaurin-gamma-1A [Musca domestica]
Woma_00006555-RA	XP_023306033.1 trypsin eta-like [Lucilia cuprina]
Woma_00006547-RA	KNC29657.1 putative RNA-directed DNA polymerase from transposon X-element [Lucilia cuprina]
Woma_00006651-RA	XP_023298792.1 cytochrome P450 6a9-like [Lucilia cuprina]
Woma_00006650-RA	AID61394.1 cytochrome P450 [Calliphora stygia]
Woma_00006318-RA	XP_023305303.1 mucin-2 [Lucilia cuprina]
Woma_00006361-RB	KNC27255.1 Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 [Lucilia cuprina]
Woma_00006360-RA	KNC27255.1 Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 [Lucilia cuprina]
Woma_00006824-RA	XP_013101123.1 PREDICTED: box A-binding factor isoform X1 [Stomoxys calcitrans]
Woma_00006820-RA	XP_023307975.1 angiopoietin-related protein 2-like [Lucilia cuprina]
Woma_00006783-RA	XP_023303587.1 zinc finger protein 502-like [Lucilia cuprina]
Woma_00006781-RA	XP_023303578.1 zinc finger protein 317-like [Lucilia cuprina]

Woma_00006780-RA	XP_023303578.1 zinc finger protein 317-like [Lucilia cuprina]
Woma_00006784-RA	XP_023303587.1 zinc finger protein 502-like [Lucilia cuprina]
Woma_00006806-RA	XP_008545617.1 PREDICTED: histone H2A-like [Microplitis demolitor]
Woma_00006819-RA	XP_023300613.1 homeobox protein unc-4 homolog [Lucilia cuprina]
Woma_00006814-RB	XP_023297889.1 alpha-protein kinase 1-like [Lucilia cuprina]
Woma_00006850-RA	XP_023307796.1 ADP-ribosylation factor 1-like [Lucilia cuprina]
Woma_00006778-RA	XP_004527474.1 zinc finger protein ZFMSA12A [Ceratitis capitata]
Woma_00006750-RB	KNC26959.1 putative gustatory receptor 98b [Lucilia cuprina]
Woma_00006752-RA	AAB34768.1 chemical-sense-related lipophilic-ligand-binding protein [Phormia regina]
Woma_00006753-RA	XP_023299671.1 general odorant-binding protein 19d-like [Lucilia cuprina]
Woma_00008655-RA	XP_014093144.1 PREDICTED: 5-hydroxytryptamine receptor 2A [Bactrocera oleae]
Woma_00008654-RA	XP_023307101.1 5-hydroxytryptamine receptor 2A-like, partial [Lucilia cuprina]
Woma_00008651-RA	KMQ87671.1 reverse transcriptase [Lasius niger]
Woma_00008708-RA	XP_023299918.1 GDNF-inducible zinc finger protein 1-like [Lucilia cuprina]
Woma_00008660-RB	KNC20935.1 putative cytochrome P450 313a4, partial [Lucilia cuprina]
Woma_00008666-RA	XP_023304687.1 dopamine N-acetyltransferase-like [Lucilia cuprina]
Woma_00008667-RA	XP_023304687.1 dopamine N-acetyltransferase-like [Lucilia cuprina]
Woma_00008681-RA	XP_023308339.1 laccase-21, partial [Lucilia cuprina]
Woma_00006960-RA	KNC32378.1 Echinoderm microtubule-associated protein-like protein [Lucilia cuprina]
Woma_00006986-RA	XP_023297623.1 cyclin-dependent kinase 12 [Lucilia cuprina]
Woma_00007029-RA	XP_017137359.1 PREDICTED: putative glycine-rich cell wall structural protein 1 [Drosophila miranda]
Woma_00007055-RC	XP_023293382.1 protein I'm not dead yet [Lucilia cuprina]
Woma_00007056-RA	XP_023293382.1 protein I'm not dead yet [Lucilia cuprina]
Woma_00007116-RA	KNC34993.1 putative phosphatidate phosphatase [Lucilia cuprina]
Woma_00007084-RA	XP_023303625.1 whirlin-like [Lucilia cuprina]
Woma_00007148-RA	XP_023306372.1 nucleolar and coiled-body phosphoprotein 1 isoform X2 [Lucilia cuprina]
Woma_00007121-RA	XP_011293135.1 PREDICTED: protein kinase 4-like, partial [Musca domestica]
Woma_00007178-RA	XP_023294359.1 zinc finger protein 782-like [Lucilia cuprina]
Woma_00007201-RA	XP_023297288.1 neurogenic protein mastermind-like, partial [Lucilia cuprina]
Woma_00007194-RA	KMQ83311.1 transposable element tc3 transposase [Lasius niger]
Woma_00007161-RA	XP_023294382.1 RNA-binding protein fusilli isoform X2 [Lucilia cuprina]
Woma_00007187-RA	XP_023299084.1 microtubule-actin cross-linking factor 1 isoform X1 [Lucilia cuprina]
Woma_00007269-RA	XP_023308229.1 pollen-specific leucine-rich repeat extensin-like protein 3 [Lucilia cuprina]
Woma_00007583-RA	XP_023295226.1 mediator of RNA polymerase II transcription subunit 30 [Lucilia cuprina]
Woma_00007677-RA	KNC21829.1 putative odorant receptor 45a [Lucilia cuprina]
Woma_00007567-RA	XP_023307979.1 mitochondrial thiamine pyrophosphate carrier-like [Lucilia cuprina]
Woma_00007599-RA	KNC33320.1 putative odorant receptor 7a [Lucilia cuprina]
Woma_00007590-RA	XP_023303348.1 apolipoprotein D-like [Lucilia cuprina]
Woma_00007669-RA	AID61317.1 odorant binding protein, partial [Calliphora stygia]
Woma_00007668-RA	AID61317.1 odorant binding protein, partial [Calliphora stygia]
Woma_00007667-RA	AID61317.1 odorant binding protein, partial [Calliphora stygia]
Woma_00007666-RA	AID61317.1 odorant binding protein, partial [Calliphora stygia]
Woma_00007664-RA	AID61317.1 odorant binding protein, partial [Calliphora stygia]
Woma_00007535-RA	XP_023295230.1 dr1-associated corepressor homolog [Lucilia cuprina]

Woma_00007556-RA	BAS69459.1 odorant-binding protein 17 [Delia platura]
Woma_00007577-RA	AID61213.1 odorant receptor, partial [Calliphora stygia]
Woma_00007696-RA	XP_023305498.1 zinc finger protein 91-like [Lucilia cuprina]
Woma_00007695-RA	XP_023305498.1 zinc finger protein 91-like [Lucilia cuprina]
Woma_00007739-RA	XP_023297241.1 synaptic vesicle glycoprotein 2A-like [Lucilia cuprina]
Woma_00007710-RA	XP_023296818.1 potassium channel subfamily K member 1-like isoform X1 [Lucilia cuprina]
Woma_00007773-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00007740-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00007741-RA	XP_023297243.1 synaptic vesicle glycoprotein 2A-like [Lucilia cuprina]
Woma_00005142-RA	XP_023305188.1 membrane metallo-endopeptidase-like 1 [Lucilia cuprina]
Woma_00005146-RA	XP_023305177.1 neprilysin-11-like [Lucilia cuprina]
Woma_00005093-RB	XP_023300882.1 alpha-tocopherol transfer protein-like [Lucilia cuprina]
Woma_00005034-RA	XP_023299996.1 trimethyllysine dioxygenase, mitochondrial-like, partial [Lucilia cuprina]
Woma_00005054-RA	XP_005187543.1 PREDICTED: discoidin domain-containing receptor 2, partial [Musca domestica]
Woma_00005096-RA	XP_023300872.1 general transcriptional corepressor trfA-like [Lucilia cuprina]
Woma_00005151-RC	XP_023305184.1 protein I'm not dead yet-like [Lucilia cuprina]
Woma_00005041-RA	XP_023306864.1 mucin-2-like [Lucilia cuprina]
Woma_00005042-RB	XP_023306866.1 keratin-associated protein 10-6-like [Lucilia cuprina]
Woma_00005842-RA	KNC28737.1 Protein dachsous [Lucilia cuprina]
Woma_00005814-RA	BAA28264.1 20kDa lectin [Sarcophaga peregrina]
Woma_00005766-RA	XP_023303879.1 paramyosin [Lucilia cuprina]
Woma_00013118-RA	XP_023292863.1 probable 4-coumarate--CoA ligase 3 [Lucilia cuprina]
Woma_00005301-RA	XP_023295312.1 mucin-5AC-like [Lucilia cuprina]
Woma_00005233-RA	XP_023300392.1 glutathione S-transferase 1-like [Lucilia cuprina]
Woma_00007506-RA	XP_023300173.1 fibrinogen C domain-containing protein 1-like [Lucilia cuprina]
Woma_00007507-RA	XP_023300173.1 fibrinogen C domain-containing protein 1-like [Lucilia cuprina]
Woma_00007504-RA	XP_023300173.1 fibrinogen C domain-containing protein 1-like [Lucilia cuprina]
Woma_00007505-RA	XP_023300173.1 fibrinogen C domain-containing protein 1-like [Lucilia cuprina]
Woma_00007503-RA	XP_023300160.1 ryncolin-1-like [Lucilia cuprina]
Woma_00007508-RA	XP_023300170.1 fibrinogen C domain-containing protein 1-like [Lucilia cuprina]
Woma_00007520-RA	XP_023307376.1 aquaporin AQPAe.a-like [Lucilia cuprina]
Woma_00007521-RA	XP_023308118.1 aquaporin AQPcic-like [Lucilia cuprina]
Woma_00007451-RA	XP_013098053.1 PREDICTED: trithorax group protein osa isoform X1 [Stomoxys calcitrans]
Woma_00007433-RA	XP_023294886.1 splicing factor, proline- and glutamine-rich-like [Lucilia cuprina]
Woma_00007434-RA	AAS13459.1 putative reverse transcriptase [Drosophila simulans]
Woma_00007384-RA	XP_023293524.1 serendipity locus protein delta-like [Lucilia cuprina]
Woma_00007386-RA	XP_023291802.1 serendipity locus protein delta-like [Lucilia cuprina]
Woma_00007510-RB	XP_023300160.1 ryncolin-1-like [Lucilia cuprina]
Woma_00007513-RA	XP_023300176.1 BRCA1-associated protein-like [Lucilia cuprina]
Woma_00007439-RA	XP_023300659.1 rho GTPase-activating protein gacZ-like [Lucilia cuprina]
Woma_00007420-RA	OXA40618.1 Transposable element Tcb1 transposase [Folsomia candida]
Woma_00007422-RA	XP_023301657.1 venom allergen 3-like [Lucilia cuprina]
Woma_00007406-RA	XP_014096438.1 PREDICTED: zinc finger protein 704-like [Bactrocera oleae]
Woma_00005499-RA	KMQ89555.1 reverse transcriptase [Lasius niger]

Woma_00005639-RA	XP_023308840.1 carbonic anhydrase 2-like [Lucilia cuprina]
Woma_00005637-RC	XP_023308849.1 carbonic anhydrase 13-like [Lucilia cuprina]
Woma_00005668-RA	XP_023309185.1 tissue factor pathway inhibitor-like [Lucilia cuprina]
Woma_00005529-RC	XP_023308779.1 protein slender lobes [Lucilia cuprina]
Woma_00006290-RA	XP_023297508.1 keratin-associated protein 6-2 [Lucilia cuprina]
Woma_00006288-RA	XP_023297507.1 BCL-6 corepressor-like protein 1 [Lucilia cuprina]
Woma_00006289-RA	XP_023297507.1 BCL-6 corepressor-like protein 1 [Lucilia cuprina]
Woma_00006195-RA	XP_023302286.1 collagen alpha-1(I) chain-like isoform X2 [Lucilia cuprina]
Woma_00005866-RA	KNC32871.1 putative gustatory receptor 98b [Lucilia cuprina]
Woma_00005861-RA	KNC32871.1 putative gustatory receptor 98b [Lucilia cuprina]
Woma_00005857-RB	XP_023298656.1 putative gustatory receptor 98b [Lucilia cuprina]
Woma_00005855-RA	PSN51582.1 putative RNA-directed DNA polymerase from transposon X-element [Blattella germanica]
Woma_00005872-RB	XP_023293485.1 host cell factor-like [Lucilia cuprina]
Woma_00008148-RA	PNF42959.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00008220-RB	XP_023291595.1 zinc finger protein 652-A-like [Lucilia cuprina]
Woma_00008114-RA	XP_023296157.1 zinc finger protein 91-like [Lucilia cuprina]
Woma_00008219-RA	XP_023291628.1 zinc finger protein 699-like [Lucilia cuprina]
Woma_00006163-RA	KNC23960.1 Coiled-coil domain-containing protein lobo [Lucilia cuprina]
Woma_00014875-RA	XP_023309497.1 pleckstrin homology-like domain family B member 1 [Lucilia cuprina]
Woma_00014950-RA	KNC26064.1 G2/mitotic-specific cyclin-A [Lucilia cuprina]
Woma_00014968-RA	XP_023297400.1 putative gustatory receptor 58c [Lucilia cuprina]
Woma_00014962-RA	AAA29987.1 sarcotoxin IIA [Sarcophaga peregrina]
Woma_00014963-RA	AAA29987.1 sarcotoxin IIA [Sarcophaga peregrina]
Woma_00014964-RA	AAA29987.1 sarcotoxin IIA [Sarcophaga peregrina]
Woma_00014965-RA	XP_023293356.1 sarcotoxin-2A-like [Lucilia cuprina]
Woma_00015013-RA	PIK42074.1 Transposon Ty3-I Gag-Pol polyprotein [Apostichopus japonicus]
Woma_00015012-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00014989-RA	DAA65009.1 TPA_exp: gag-pol protein [Drosophila simulans]
Woma_00015109-RA	KRZ65659.1 Retrovirus-related Pol polyprotein from transposon TNT 1-94 [Trichinella papuae]
Woma_00015157-RA	PNF35856.1 putative RNA-directed DNA polymerase from transposon X-element, partial [Cryptotermes secundus]
Woma_00015153-RA	DAA65006.1 TPA_exp: gag-pol protein [Drosophila sechellia]
Woma_00015221-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00015235-RB	PIK49151.1 putative Gag-pol polyprotein [Apostichopus japonicus]
Woma_00015290-RA	PIK50756.1 Retrovirus-related Pol polyprotein from transposon [Apostichopus japonicus]
Woma_00015323-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016197-RA	KMQ93693.1 reverse transcriptase [Lasius niger]
Woma_00015434-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00015363-RB	XP_023293160.1 high affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8-like [Lucilia cuprina]
Woma_00015388-RA	KRY26687.1 Retrovirus-related Pol polyprotein from transposon TNT 1-94 [Trichinella spiralis]
Woma_00015393-RA	KRY26687.1 Retrovirus-related Pol polyprotein from transposon TNT 1-94 [Trichinella spiralis]
Woma_00015496-RA	XP_023308270.1 RING-H2 finger protein ATL51-like, partial [Lucilia cuprina]

Woma_00015565-RA	PIK48899.1 Retrovirus-related Pol polyprotein from transposon [Apostichopus japonicus]
Woma_00015563-RA	PNF14434.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00015600-RA	XP_023297828.1 general odorant-binding protein 56a-like [Lucilia cuprina]
Woma_00015591-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00015715-RB	KMQ91779.1 integrase core domain protein [Lasius niger]
Woma_00015704-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00014003-RC	KNC20976.1 Serine protease easter [Lucilia cuprina]
Woma_00013347-RA	AAC63387.1 23kDa heat shock protein ScHSP23 [Sarcophaga crassipalpis]
Woma_00013346-RB	XP_023304388.1 heat shock protein 67B2-like [Lucilia cuprina]
Woma_00013363-RA	XP_023309203.1 probable serine hydrolase [Lucilia cuprina]
Woma_00014403-RA	XP_011290715.1 PREDICTED: phosphatidate phosphatase LPIN1 isoform X2 [Musca domestica]
Woma_00014512-RA	XP_023299237.1 pheromone-binding protein-related protein 6-like [Lucilia cuprina]
Woma_00013732-RA	KMQ88472.1 retrovirus-related gag-pol polyprotein [Lasius niger]
Woma_00013743-RA	AAL56548.1 pol polyprotein, partial [Anopheles gambiae]
Woma_00013496-RA	KMQ83311.1 transposable element tc3 transposase [Lasius niger]
Woma_00013548-RA	EFN65994.1 Retrovirus-related Pol polyprotein from transposon TNT 1-94, partial [Camponotus floridanus]
Woma_00013573-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00013674-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00013668-RB	KXJ24126.1 Retrovirus-related Pol polyprotein from transposon 297 [Exaiptasia pallida]
Woma_00013880-RA	XP_005189426.1 PREDICTED: serum response factor homolog isoform X1 [Musca domestica]
Woma_00009277-RA	DAA65008.1 TPA_exp: gag-pol protein [Drosophila yakuba]
Woma_00009299-RA	XP_023309632.1 meiotic nuclear division protein 1 homolog [Lucilia cuprina]
Woma_00009295-RA	XP_023309647.1 double-headed protease inhibitor, submandibular gland-like [Lucilia cuprina]
Woma_00009286-RA	ANN46821.1 Pol protein [Drosophila simulans]
Woma_00011317-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00011275-RA	XP_023298991.1 ATP synthase subunit b, mitochondrial-like [Lucilia cuprina]
Woma_00011272-RA	XP_023298233.1 alpha-2Db adrenergic receptor [Lucilia cuprina]
Woma_00009564-RA	KNC34166.1 Tropomyosin-1, isoforms 9A/A/B [Lucilia cuprina]
Woma_00009520-RA	XP_023292046.1 arrestin domain-containing protein 3-like [Lucilia cuprina]
Woma_00009348-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009349-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009345-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009342-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009343-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009341-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009346-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009347-RA	XP_023302771.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009344-RA	XP_023302779.1 transcription factor grauzone-like [Lucilia cuprina]
Woma_00009365-RA	KFM75221.1 Transposable element Tcb1 transposase, partial [Stegodyphus mimosarum]
Woma_00010953-RA	XP_023301757.1 transcriptional regulator ATRX-like isoform X1 [Lucilia cuprina]

Woma_00010871-RA	XP_023297396.1 retinal guanylyl cyclase 2-like [Lucilia cuprina]
Woma_00010868-RA	KMQ91018.1 reverse transcriptase [Lasius niger]
Woma_00010865-RA	KNC22213.1 Down syndrome cell adhesion molecule-like protein Dscam2, partial [Lucilia cuprina]
Woma_00010922-RA	XP_013099300.1 PREDICTED: calcineurin-binding protein cabin-1-like [Stomoxys calcitrans]
Woma_00010928-RA	XP_023303956.1 toll-interacting protein-like [Lucilia cuprina]
Woma_00009469-RA	XP_023299934.1 neutral and basic amino acid transport protein rBAT-like isoform X1 [Lucilia cuprina]
Woma_00009464-RA	XP_023298391.1 active regulator of SIRT1-like [Lucilia cuprina]
Woma_00009462-RB	XP_023298389.1 tetraspanin-6-like [Lucilia cuprina]
Woma_00009640-RB	XP_023292507.1 flocculation protein FLO11-like [Lucilia cuprina]
Woma_00009610-RA	PNF17579.1 putative RNA-directed DNA polymerase from transposon X-element, partial [Cryptotermes secundus]
Woma_00009613-RA	CAP09075.1 minos transposase [Drosophila hydei]
Woma_00009639-RA	XP_023294019.1 rho guanine nucleotide exchange factor 11-like [Lucilia cuprina]
Woma_00009831-RA	XP_014614959.1 PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like, partial [Polistes canadensis]
Woma_00009809-RA	XP_015438116.1 PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like [Dufourea novaeangliae]
Woma_00009844-RA	KMQ83311.1 transposable element tc3 transposase [Lasius niger]
Woma_00009740-RA	XP_023297714.1 cell division cycle and apoptosis regulator protein 1-like isoform X2 [Lucilia cuprina]
Woma_00009739-RA	XP_023297713.1 cell division cycle and apoptosis regulator protein 1-like isoform X1 [Lucilia cuprina]
Woma_00009773-RB	XP_023291465.1 protein ELYS homolog [Lucilia cuprina]
Woma_00009700-RA	XP_023293290.1 RING finger protein 219-like [Lucilia cuprina]
Woma_00009768-RA	XP_023291481.1 zinc finger protein 271-like isoform X2 [Lucilia cuprina]
Woma_00009769-RA	XP_023291480.1 zinc finger protein 11-like isoform X1 [Lucilia cuprina]
Woma_00009899-RA	XP_023303161.1 zinc finger protein 287-like [Lucilia cuprina]
Woma_00009923-RA	XP_023295890.1 INO80 complex subunit D-like [Lucilia cuprina]
Woma_00009926-RA	XP_023295891.1 farnesol dehydrogenase-like [Lucilia cuprina]
Woma_00009925-RA	XP_023295893.1 farnesol dehydrogenase-like [Lucilia cuprina]
Woma_00010052-RA	XP_023298844.1 neuropeptide-like 3 [Lucilia cuprina]
Woma_00010020-RA	XP_023301348.1 probable ubiquitin-conjugating enzyme E2 31 [Lucilia cuprina]
Woma_00009985-RA	AGE31793.1 transformer [Cochliomyia hominivorax]
Woma_00009986-RA	ACS34687.2 transformer [Lucilia cuprina]
Woma_00010046-RA	XP_023307709.1 A-kinase anchor protein 14-like isoform X1 [Lucilia cuprina]
Woma_00013538-RA	XP_023305931.1 trypsin-1-like, partial [Lucilia cuprina]
Woma_00013537-RA	XP_023305931.1 trypsin-1-like, partial [Lucilia cuprina]
Woma_00013534-RA	AID61414.1 cytochrome P450 [Calliphora stygia]
Woma_00011773-RA	XP_023291044.1 odorant receptor 13a-like [Lucilia cuprina]
Woma_00010253-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00010218-RA	XP_023302843.1 peptidoglycan-recognition protein 3 isoform X2 [Lucilia cuprina]
Woma_00010185-RA	XP_023301443.1 sensory neuron membrane protein 2-like [Lucilia cuprina]
Woma_00010308-RA	KNC29657.1 putative RNA-directed DNA polymerase from transposon X-element [Lucilia cuprina]
Woma_00010367-RA	XP_023308382.1 RYamide receptor-like [Lucilia cuprina]
Woma_00010372-RA	XP_011563256.1 PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like [Plutella xylostella]
Woma_00010422-RA	KHJ88486.1 transposase [Oesophagostomum dentatum]
Woma_00010493-RA	XP_023301201.1 protein scylla-like [Lucilia cuprina]
Woma_00010482-RA	KNC28353.1 Protein scylla, partial [Lucilia cuprina]

Woma_00010488-RA	XP_023303114.1 protein scylla-like [Lucilia cuprina]
Woma_00010661-RA	XP_013099303.1 PREDICTED: dynein heavy chain 7, axonemal [Stomoxys calcitrans]
Woma_00010638-RA	XP_023299360.1 nitric oxide synthase [Lucilia cuprina]
Woma_00010760-RA	XP_023294530.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010761-RB	XP_023304878.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010766-RA	XP_023304878.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010764-RA	XP_023304878.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010763-RA	XP_023304878.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010752-RA	KHJ84698.1 putative transposase [Oesophagostomum dentatum]
Woma_00010771-RA	XP_023303910.1 apolipoprotein D-like [Lucilia cuprina]
Woma_00010759-RA	XP_023294530.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00010684-RB	XP_013114958.1 PREDICTED: transcription factor grauzone-like [Stomoxys calcitrans]
Woma_00010680-RA	XP_023294530.1 lysosomal aspartic protease-like [Lucilia cuprina]
Woma_00011334-RA	XP_023304701.1 odorant receptor Or2-like [Lucilia cuprina]
Woma_00011380-RA	XP_023299313.1 putative mediator of RNA polymerase II transcription subunit 26 [Lucilia cuprina]
Woma_00011340-RA	XP_023304803.1 electron transfer flavoprotein beta subunit lysine methyltransferase-like [Lucilia cuprina]
Woma_00013416-RA	XP_023307356.1 low density lipoprotein receptor adapter protein 1-B-like [Lucilia cuprina]
Woma_00010838-RA	KNC23576.1 Small conductance calcium-activated potassium channel protein [Lucilia cuprina]
Woma_00010792-RA	XP_023295233.1 enhancer of split mgamma protein-like [Lucilia cuprina]
Woma_00010815-RA	XP_023295224.1 zinc finger protein 1 homolog [Lucilia cuprina]
Woma_00010853-RA	XP_023293933.1 zinc finger protein 184-like [Lucilia cuprina]
Woma_00010785-RA	XP_023292191.1 alpha-protein kinase 1-like [Lucilia cuprina]
Woma_00010801-RB	XP_023295261.1 venom peptide CtAPI-like [Lucilia cuprina]
Woma_00010807-RA	XP_023295225.1 zinc finger protein 1 homolog [Lucilia cuprina]
Woma_00010845-RA	XP_023300060.1 golgin subfamily A member 6-like protein 6 [Lucilia cuprina]
Woma_00014260-RA	XP_023291366.1 zinc finger and SCAN domain-containing protein 12 [Lucilia cuprina]
Woma_00014246-RA	XP_023303417.1 fibrinogen-like protein 1 [Lucilia cuprina]
Woma_00014250-RA	XP_005180483.1 PREDICTED: tyrosine-protein phosphatase corkscrew isoform X3 [Musca domestica]
Woma_00014284-RA	XP_014091565.1 PREDICTED: zinc finger and BTB domain-containing protein 24-like [Bactrocera oleae]
Woma_00014292-RA	XP_023291420.1 odorant receptor 42b-like [Lucilia cuprina]
Woma_00014293-RA	AID61230.1 odorant receptor [Calliphora stygia]
Woma_00014283-RA	XP_023291380.1 gastrula zinc finger protein xFG20-1-like [Lucilia cuprina]
Woma_00014264-RA	XP_023291392.1 zinc finger protein 107-like isoform X1 [Lucilia cuprina]
Woma_00012759-RA	XP_023301232.1 probable basic-leucine zipper transcription factor S [Lucilia cuprina]
Woma_00012795-RA	XP_023307679.1 amine sulfotransferase-like [Lucilia cuprina]
Woma_00012797-RA	XP_023307677.1 sulfotransferase 1 family member D1-like [Lucilia cuprina]
Woma_00012744-RA	XP_023299573.1 peritrophin-44-like [Lucilia cuprina]
Woma_00012799-RB	XP_023307677.1 sulfotransferase 1 family member D1-like [Lucilia cuprina]
Woma_00012798-RB	KNC34646.1 Polypeptide N-acetylgalactosaminyltransferase 8 [Lucilia cuprina]
Woma_00012801-RA	KNC34646.1 Polypeptide N-acetylgalactosaminyltransferase 8 [Lucilia cuprina]
Woma_00012800-RA	KNC34646.1 Polypeptide N-acetylgalactosaminyltransferase 8 [Lucilia cuprina]

Woma_00012486-RA	XP_023292299.1 ras-related protein Rab-26 [Lucilia cuprina]
Woma_00012475-RB	XP_023304802.1 lectin subunit alpha-like [Lucilia cuprina]
Woma_00013007-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00011100-RA	XP_023294134.1 sex peptide receptor-like [Lucilia cuprina]
Woma_00011102-RB	XP_005181379.1 PREDICTED: apolipoprotein D [Musca domestica]
Woma_00007839-RA	XP_023304152.1 zinc finger protein 425-like [Lucilia cuprina]
Woma_00007811-RA	AID61460.1 cytochrome P450, partial [Calliphora stygia]
Woma_00007813-RA	AID61460.1 cytochrome P450, partial [Calliphora stygia]
Woma_00007886-RA	XP_023294027.1 malate dehydrogenase, mitochondrial-like [Lucilia cuprina]
Woma_00007885-RA	XP_023294028.1 malate dehydrogenase, mitochondrial-like [Lucilia cuprina]
Woma_00007781-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00007902-RA	XP_023292043.1 mucin-2-like [Lucilia cuprina]
Woma_00007972-RA	XP_023304134.1 tensin-1-like [Lucilia cuprina]
Woma_00008053-RB	AID61247.1 odorant receptor, partial [Calliphora stygia]
Woma_00007994-RA	XP_023296091.1 protein phosphatase 1 regulatory subunit 12A isoform X9 [Lucilia cuprina]
Woma_00008054-RA	AID61247.1 odorant receptor, partial [Calliphora stygia]
Woma_00007979-RA	XP_019893896.1 PREDICTED: basic-leucine zipper transcription factor A-like, partial [Musca domestica]
Woma_00008016-RA	XP_023308126.1 chorion protein S18 [Lucilia cuprina]
Woma_00009035-RA	KNC29869.1 Protein bric-a-brac 1 [Lucilia cuprina]
Woma_00009096-RB	XP_023309003.1 DNA-binding protein D-ETS-3-like [Lucilia cuprina]
Woma_00009149-RA	XP_023307895.1 serine protease 1-like [Lucilia cuprina]
Woma_00009082-RB	KNC23496.1 Protein still life, isoform SIF type 1 [Lucilia cuprina]
Woma_00008398-RA	KNC22252.1 putative gustatory receptor 66a [Lucilia cuprina]
Woma_00008396-RA	XP_023299989.1 lectin subunit alpha-like [Lucilia cuprina]
Woma_00008350-RB	XP_023297755.1 otoferlin-like [Lucilia cuprina]
Woma_00008402-RA	XP_023292787.1 glycine-rich cell wall structural protein 1.8-like [Lucilia cuprina]
Woma_00008267-RA	XP_023301508.1 shugoshin [Lucilia cuprina]
Woma_00008531-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00008609-RA	XP_013111937.1 PREDICTED: capon-like protein, partial [Stomoxys calcitrans]
Woma_00008569-RA	XP_023306132.1 pupal cuticle protein Edg-78E-like [Lucilia cuprina]
Woma_00008528-RB	XP_023292722.1 protein rhomboid-like [Lucilia cuprina]
Woma_00008723-RA	AAA29354.1 reverse transcriptase-like protein, partial [Aedes aegypti]
Woma_00008735-RA	KNC34952.1 putative G-protein coupled receptor Mth-like 11 [Lucilia cuprina]
Woma_00008732-RA	KNC34952.1 putative G-protein coupled receptor Mth-like 11 [Lucilia cuprina]
Woma_00008776-RA	ADI87385.1 putative chitin binding protein [Lucilia sericata]
Woma_00012524-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00008836-RA	XP_023304418.1 enoyl-CoA delta isomerase 2, mitochondrial-like [Lucilia cuprina]
Woma_00008846-RA	XP_023304408.1 A-agglutinin anchorage subunit-like [Lucilia cuprina]
Woma_00008813-RA	XP_023300275.1 lectin subunit alpha-like [Lucilia cuprina]
Woma_00008918-RA	KNC32871.1 putative gustatory receptor 98b [Lucilia cuprina]
Woma_00008911-RA	XP_013107470.1 PREDICTED: DNA-directed RNA polymerase III subunit RPC1 [Stomoxys calcitrans]
Woma_00009002-RA	KNC28259.1 Coiled-coil domain-containing protein lobo [Lucilia cuprina]
Woma_00009237-RA	PNF14434.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]

Woma_00009265-RB	PNF17579.1 putative RNA-directed DNA polymerase from transposon X-element, partial [Cryptotermes secundus]
Woma_00009233-RA	PNF14434.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00001773-RA	XP_023300606.1 alcohol dehydrogenase-like [Lucilia cuprina]
Woma_00001714-RA	XP_023299240.1 ribonucleoprotein RB97D isoform X2 [Lucilia cuprina]
Woma_00001802-RA	XP_013100134.1 PREDICTED: protein timeless homolog [Stomoxys calcitrans]
Woma_00001757-RA	XP_023292069.1 calcyphosin-like isoform X1 [Lucilia cuprina]
Woma_00001806-RB	XP_023297488.1 integumentary mucin C.1-like [Lucilia cuprina]
Woma_00002188-RA	XP_023291750.1 cAMP-specific 3',5'-cyclic phosphodiesterase-like [Lucilia cuprina]
Woma_00002191-RA	XP_023291757.1 cAMP-specific 3',5'-cyclic phosphodiesterase-like [Lucilia cuprina]
Woma_00001973-RA	KNC28397.1 Trypsin eta, partial [Lucilia cuprina]
Woma_00001959-RA	XP_023295920.1 transferrin-like [Lucilia cuprina]
Woma_00001939-RA	XP_023299038.1 myosin-2 heavy chain-like [Lucilia cuprina]
Woma_00001937-RA	KMQ88467.1 tc1-like transposase protein [Lasius niger]
Woma_00002175-RA	XP_023291224.1 lectin subunit alpha-like [Lucilia cuprina]
Woma_00001914-RA	XP_023291826.1 pro-resilin-like [Lucilia cuprina]
Woma_00002258-RA	XP_023306429.1 zinc finger protein 2-like [Lucilia cuprina]
Woma_00002251-RA	XP_013108764.1 PREDICTED: E3 ubiquitin-protein ligase UBR3 [Stomoxys calcitrans]
Woma_00002323-RA	XP_023298004.1 probable 4-coumarate--CoA ligase 3 [Lucilia cuprina]
Woma_00002429-RA	XP_023304947.1 probable 4-coumarate--CoA ligase 3 [Lucilia cuprina]
Woma_00002299-RA	XP_013114958.1 PREDICTED: transcription factor grauzone-like [Stomoxys calcitrans]
Woma_00002298-RA	XP_013114958.1 PREDICTED: transcription factor grauzone-like [Stomoxys calcitrans]
Woma_00002297-RA	XP_023298129.1 zinc finger protein 208-like [Lucilia cuprina]
Woma_00002296-RA	XP_023298129.1 zinc finger protein 208-like [Lucilia cuprina]
Woma_00002308-RA	XP_023298094.1 probable 4-coumarate--CoA ligase 1 isoform X1 [Lucilia cuprina]
Woma_00002303-RA	XP_023298127.1 zinc finger protein 729-like [Lucilia cuprina]
Woma_00002302-RA	XP_023298127.1 zinc finger protein 729-like [Lucilia cuprina]
Woma_00002301-RA	XP_023298127.1 zinc finger protein 729-like [Lucilia cuprina]
Woma_00002300-RA	XP_013114958.1 PREDICTED: transcription factor grauzone-like [Stomoxys calcitrans]
Woma_00002305-RA	XP_023298126.1 zinc finger protein 208-like [Lucilia cuprina]
Woma_00002304-RA	XP_023298127.1 zinc finger protein 729-like [Lucilia cuprina]
Woma_00002324-RA	XP_023298004.1 probable 4-coumarate--CoA ligase 3 [Lucilia cuprina]
Woma_00002381-RA	XP_013098329.1 PREDICTED: protein sprint isoform X1 [Stomoxys calcitrans]
Woma_00002224-RA	AID61355.1 esterase, partial [Calliphora stygia]
Woma_00002565-RA	XP_023294308.1 flocculation protein FLO11-like [Lucilia cuprina]
Woma_00002582-RA	XP_023294797.1 broad-complex core protein isoforms 1/2/3/4/5 [Lucilia cuprina]
Woma_00002576-RA	XP_023294310.1 signal transducer and activator of transcription C-like [Lucilia cuprina]
Woma_00002607-RA	XP_023293962.1 centrosomal and chromosomal factor-like [Lucilia cuprina]
Woma_00002526-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00002572-RA	XP_023294311.1 collagen alpha-5(IV) chain [Lucilia cuprina]
Woma_00002476-RA	XP_023296780.1 salivary glue protein Sgs-3-like [Lucilia cuprina]
Woma_00002834-RA	XP_023306842.1 zinc finger protein 226-like [Lucilia cuprina]
Woma_00002859-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]

Woma_00002835-RA	XP_023307829.1 zinc finger protein 540-like [Lucilia cuprina]
Woma_00002868-RA	XP_023293521.1 ryncolin-4-like [Lucilia cuprina]
Woma_00002824-RA	XP_023307829.1 zinc finger protein 540-like [Lucilia cuprina]
Woma_00002826-RA	XP_023307829.1 zinc finger protein 540-like [Lucilia cuprina]
Woma_00002752-RA	KMQ87671.1 reverse transcriptase [Lasius niger]
Woma_00002863-RA	AID61460.1 cytochrome P450, partial [Calliphora stygia]
Woma_00002763-RB	XP_005178160.1 PREDICTED: protein doublesex isoform X2 [Musca domestica]
Woma_00002864-RA	XP_023304796.1 saccharopine dehydrogenase-like oxidoreductase [Lucilia cuprina]
Woma_00002862-RA	XP_023292858.1 4-coumarate--CoA ligase-like 7 [Lucilia cuprina]
Woma_00003076-RA	XP_023300216.1 UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like protein 1 [Lucilia cuprina]
Woma_00003014-RA	XP_023302134.1 sodium channel protein Nach-like [Lucilia cuprina]
Woma_00002969-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00002970-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00002967-RA	XP_023160382.1 serine protease 1 [Drosophila hydei]
Woma_00002973-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00002933-RA	XP_023296535.1 zinc finger protein 43-like [Lucilia cuprina]
Woma_00003127-RA	XP_023304035.1 proteoglycan 4-like [Lucilia cuprina]
Woma_00002968-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00002972-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00003134-RA	KNC30428.1 Protein sickie [Lucilia cuprina]
Woma_00002917-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00002971-RA	XP_023293217.1 serine protease 1-like [Lucilia cuprina]
Woma_00003320-RA	XP_023293284.1 very low-density lipoprotein receptor-like [Lucilia cuprina]
Woma_00003455-RB	XP_023293684.1 UDP-glucuronosyltransferase-like [Lucilia cuprina]
Woma_00003360-RA	OXA39388.1 Transposable element Tc3 transposase [Folsomia candida]
Woma_00003369-RB	XP_013098269.1 PREDICTED: maternal protein pumilio isoform X1 [Stomoxys calcitrans]
Woma_00003384-RA	XP_023305913.1 zinc finger protein 235-like [Lucilia cuprina]
Woma_00003383-RA	XP_023305913.1 zinc finger protein 235-like [Lucilia cuprina]
Woma_00015759-RA	AID61262.1 gustatory receptor [Calliphora stygia]
Woma_00015819-RA	PFX21553.1 Retrovirus-related Pol polyprotein [Stylophora pistillata]
Woma_00015932-RA	KNC28259.1 Coiled-coil domain-containing protein lobo [Lucilia cuprina]
Woma_00015929-RA	KRZ84108.1 Retrovirus-related Pol polyprotein from transposon TNT 1-94 [Trichinella sp. T8]
Woma_00015922-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00015923-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00015903-RA	KMQ83299.1 reverse transcriptase [Lasius niger]
Woma_00015970-RA	XP_017092979.1 PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like [Drosophila bipectinata]
Woma_00016007-RA	XP_011209308.1 PREDICTED: armadillo repeat-containing protein 8 isoform X1 [Bactrocera dorsalis]
Woma_00016040-RA	KMQ88472.1 retrovirus-related gag-pol polyprotein [Lasius niger]
Woma_00016069-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00016068-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00016084-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016158-RA	KMQ85192.1 reverse transcriptase [Lasius niger]
Woma_00016153-RA	KNC28259.1 Coiled-coil domain-containing protein lobo [Lucilia cuprina]

Woma_00016265-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016289-RA	PNF25410.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00016420-RB	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016489-RB	EFA07205.2 Retrovirus-related Pol polyprotein from transposon 297-like Protein [Tribolium castaneum]
Woma_00016492-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00016476-RA	XP_023292418.1 phosphate-regulating neutral endopeptidase-like [Lucilia cuprina]
Woma_00016497-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00016469-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016507-RA	XP_023295384.1 mucin-22-like [Lucilia cuprina]
Woma_00016537-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016573-RA	XP_023300376.1 RING-H2 finger protein ATL60-like, partial [Lucilia cuprina]
Woma_00016654-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016655-RA	KFM59232.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00016657-RA	XP_015509714.1 PREDICTED: putative 115 kDa protein in type-1 retrotransposable element R1DM [Neodiprion lecontei]
Woma_00016669-RA	XP_023300376.1 RING-H2 finger protein ATL60-like, partial [Lucilia cuprina]
Woma_00016691-RA	KMQ92168.1 retrovirus-related pol polyprotein from transposon tnt 1-94 [Lasius niger]
Woma_00000056-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000147-RA	XP_023292674.1 serine/threonine-protein kinase S6KL [Lucilia cuprina]
Woma_00000148-RA	XP_023292675.1 protein bangles and beads [Lucilia cuprina]
Woma_00000250-RA	KNC23846.1 putative odorant receptor 94b, partial [Lucilia cuprina]
Woma_00000060-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000061-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000062-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000063-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000088-RA	XP_023291822.1 UV excision repair protein RAD23 homolog B-like [Lucilia cuprina]
Woma_00000078-RA	XP_023291814.1 E3 ubiquitin-protein ligase highwire [Lucilia cuprina]
Woma_00000197-RA	XP_023298405.1 RING finger protein 141-like isoform X1 [Lucilia cuprina]
Woma_00000113-RA	XP_023296032.1 glutamic acid-rich protein [Lucilia cuprina]
Woma_00000057-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000058-RA	XP_023297950.1 zinc finger protein 271-like [Lucilia cuprina]
Woma_00000249-RA	KNC23846.1 putative odorant receptor 94b, partial [Lucilia cuprina]
Woma_00000224-RA	XP_023298540.1 voltage-dependent calcium channel type A subunit alpha-1-like [Lucilia cuprina]
Woma_00000358-RA	KNC31677.1 Titin, partial [Lucilia cuprina]
Woma_00000548-RA	XP_023308306.1 venom serine protease Bi-VSP-like [Lucilia cuprina]
Woma_00000468-RA	XP_005187755.1 PREDICTED: serine proteases 1/2 [Musca domestica]
Woma_00000469-RB	XP_023307896.1 serine protease 1-like [Lucilia cuprina]
Woma_00000466-RA	XP_011293756.1 PREDICTED: serine proteases 1/2-like [Musca domestica]
Woma_00000486-RA	XP_013111603.1 PREDICTED: putative carbonic anhydrase 3 [Stomoxys calcitrans]

Woma_00000677-RA	XP_023297062.1 neprilysin-2-like [Lucilia cuprina]
Woma_00000642-RA	XP_023309101.1 glutathione S-transferase 1-1-like [Lucilia cuprina]
Woma_00000650-RA	XP_023309071.1 fatty acyl-CoA reductase wat-like [Lucilia cuprina]
Woma_00000853-RA	XP_014087565.1 PREDICTED: polymerase delta-interacting protein 3-like [Bactrocera oleae]
Woma_00000648-RA	XP_023309071.1 fatty acyl-CoA reductase wat-like [Lucilia cuprina]
Woma_00000649-RA	XP_023309071.1 fatty acyl-CoA reductase wat-like [Lucilia cuprina]
Woma_00000620-RA	XP_023308947.1 histone H1, early embryonic-like [Lucilia cuprina]
Woma_00000849-RA	XP_023293869.1 protein sidekick-2-like [Lucilia cuprina]
Woma_00001062-RA	XP_023305867.1 protein RRNAD1-like [Lucilia cuprina]
Woma_00001236-RA	XP_023294085.1 transcription initiation factor TFIID subunit 1-like [Lucilia cuprina]
Woma_00001250-RA	XP_023294084.1 protein PFC0760c-like [Lucilia cuprina]
Woma_00000974-RA	KMQ86082.1 gag-pol polyprotein [Lasius niger]
Woma_00001100-RA	XP_013100224.1 PREDICTED: helicase domino [Stomoxys calcitrans]
Woma_00001121-RA	XP_023292637.1 zinc finger protein 91-like [Lucilia cuprina]
Woma_00001124-RA	XP_023292636.1 zinc finger protein 62 homolog [Lucilia cuprina]
Woma_00001123-RA	XP_023292636.1 zinc finger protein 62 homolog [Lucilia cuprina]
Woma_00001109-RA	OXA40618.1 Transposable element Tcb1 transposase [Folsomia candida]
Woma_00001159-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00001190-RA	XP_008195242.1 PREDICTED: transposable element Tcb2 transposase [Tribolium castaneum]
Woma_00001316-RA	XP_013104928.1 PREDICTED: putative mediator of RNA polymerase II transcription subunit 29 [Stomoxys calcitrans]
Woma_00001282-RA	XP_023291356.1 actin, gamma-like [Lucilia cuprina]
Woma_00001281-RA	XP_023291339.1 actin-like [Lucilia cuprina]
Woma_00001280-RA	XP_023291339.1 actin-like [Lucilia cuprina]
Woma_00001457-RA	XP_023297878.1 muscle M-line assembly protein unc-89-like isoform X3 [Lucilia cuprina]
Woma_00001335-RA	XP_023303911.1 homeobox protein invected [Lucilia cuprina]
Woma_00001334-RA	PNF14434.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00001332-RA	XP_023303720.1 chitinase-like protein Idgf5 [Lucilia cuprina]
Woma_00001441-RA	XP_023307148.1 hydroxysteroid dehydrogenase-like protein 2 [Lucilia cuprina]
Woma_00001466-RA	KMQ91018.1 reverse transcriptase [Lasius niger]
Woma_00001687-RA	KMQ83311.1 transposable element tc3 transposase [Lasius niger]
Woma_00001539-RA	PFX21553.1 Retrovirus-related Pol polyprotein [Stylophora pistillata]
Woma_00001557-RB	XP_023291437.1 cytokine receptor-like [Lucilia cuprina]
Woma_00001675-RB	XP_023296560.1 protein arginine N-methyltransferase 1-B-like [Lucilia cuprina]
Woma_00011219-RA	KNC34793.1 Homeotic protein proboscipedia [Lucilia cuprina]
Woma_00011217-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00015691-RA	XP_023294518.1 aquaporin-like [Lucilia cuprina]
Woma_00011495-RA	PNF14434.1 putative RNA-directed DNA polymerase from transposon X-element [Cryptotermes secundus]
Woma_00011539-RA	XP_023303072.1 homeobox protein caupolican-like, partial [Lucilia cuprina]
Woma_00013301-RA	XP_023302924.1 receptor expression-enhancing protein 1-like [Lucilia cuprina]
Woma_00013274-RA	XP_023295158.1 GPI ethanolamine phosphate transferase 1 [Lucilia cuprina]
Woma_00011603-RC	KNC30388.1 Protein furry, partial [Lucilia cuprina]
Woma_00011605-RB	XP_013104403.1 PREDICTED: protein furry isoform X1 [Stomoxys calcitrans]

Woma_00011696-RA	KNC23445.1 Broad-complex core protein isoforms 1/2/3/4/5 [Lucilia cuprina]
Woma_00011658-RA	XP_017475959.1 PREDICTED: putative nuclease HARBI1 isoform X1 [Rhagoletis zephyria]
Woma_00011673-RA	XP_023293734.1 ficolin-2-like [Lucilia cuprina]
Woma_00011616-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00011637-RA	ANN46821.1 Pol protein [Drosophila simulans]
Woma_00011889-RA	KNC28442.1 putative cytochrome P450 28d1 [Lucilia cuprina]
Woma_00011891-RA	XP_023294720.1 probable cytochrome P450 28d1 [Lucilia cuprina]
Woma_00011890-RA	XP_023294723.1 probable cytochrome P450 28d1 [Lucilia cuprina]
Woma_00012117-RB	XP_023297421.1 protein vav-like isoform X2 [Lucilia cuprina]
Woma_00012134-RA	XP_023295053.1 probable serine/threonine-protein kinase DDB_G0277071 [Lucilia cuprina]
Woma_00012112-RA	XP_023300973.1 putative gustatory receptor 2a [Lucilia cuprina]
Woma_00012135-RA	XP_023295045.1 bolA-like protein 3 [Lucilia cuprina]
Woma_00012236-RA	XP_023292431.1 serine protease easter-like [Lucilia cuprina]
Woma_00014552-RA	XP_023298484.1 mucin-22-like [Lucilia cuprina]
Woma_00014488-RA	XP_023297498.1 galectin-4-like [Lucilia cuprina]
Woma_00012254-RA	XP_023293713.1 facilitated trehalose transporter Tret1-like [Lucilia cuprina]
Woma_00012256-RA	XP_023293713.1 facilitated trehalose transporter Tret1-like [Lucilia cuprina]
Woma_00012306-RA	XP_023304266.1 A disintegrin and metalloproteinase with thrombospondin motifs 16 isoform X1 [Lucilia cuprina]
Woma_00012305-RA	XP_019894820.1 PREDICTED: homologous-pairing protein 2 homolog isoform X2 [Musca domestica]
Woma_00014316-RA	XP_013111068.1 PREDICTED: peritrophin-1-like [Stomoxys calcitrans]
Woma_00012314-RA	KNC31588.1 Synapsin, partial [Lucilia cuprina]
Woma_00012370-RA	XP_023301072.1 A-kinase anchor protein 14 [Lucilia cuprina]
Woma_00012430-RA	P34257.1 RecName: Full=Transposable element Tc3 transposase
Woma_00012667-RA	XP_023306970.1 endocuticle structural protein SgAbd-6-like [Lucilia cuprina]
Woma_00012678-RA	XP_023298309.1 H/ACA ribonucleoprotein complex subunit 1-like [Lucilia cuprina]
Woma_00012703-RA	XP_013118187.1 PREDICTED: eye-specific diacylglycerol kinase-like, partial [Stomoxys calcitrans]
Woma_00012704-RA	PNF20480.1 RNA-directed DNA polymerase from mobile element jockey [Cryptotermes secundus]
Woma_00012866-RB	XP_017869471.1 PREDICTED: omega-amidase NIT2-like [Drosophila arizonae]
Woma_00012865-RA	XP_023292398.1 omega-amidase NIT2-A-like [Lucilia cuprina]
Woma_00012825-RB	XP_023291423.1 cadherin-related tumor suppressor-like [Lucilia cuprina]
Woma_00014187-RA	KMQ83660.1 reverse transcriptase [Lasius niger]
Woma_00012924-RA	OXA39388.1 Transposable element Tc3 transposase [Folsomia candida]
Woma_00012928-RA	XP_023291304.1 rhomboid-related protein 3 isoform X1 [Lucilia cuprina]
Woma_00012938-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00012908-RA	XP_023306554.1 circadian clock-controlled protein-like [Lucilia cuprina]
Woma_00012899-RA	KNC27553.1 putative protein kinase C delta type, partial [Lucilia cuprina]
Woma_00013143-RA	XP_019895193.1 PREDICTED: neural-cadherin isoform X5 [Musca domestica]
Woma_00013100-RA	KFM56688.1 Retrovirus-related Pol polyprotein from transposon 412, partial [Stegodyphus mimosarum]
Woma_00013084-RA	PNF17579.1 putative RNA-directed DNA polymerase from transposon X-element, partial [Cryptotermes secundus]
Woma_00013096-RA	XP_017490395.1 PREDICTED: nucleic-acid-binding protein from mobile element jockey-like, partial [Rhagoletis zephyria]

Woma_00003308-RB	XP_023297228.1 RNA binding protein fox-1 homolog 1-like isoform X5 [Lucilia cuprina]
Woma_00003314-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00003561-RA	XP_023299610.1 P protein-like [Lucilia cuprina]
Woma_00003627-RA	XP_023306105.1 cell wall transcription factor ACE2 [Lucilia cuprina]
Woma_00003764-RA	XP_023308464.1 laccase-2-like [Lucilia cuprina]
Woma_00003927-RA	XP_023297846.1 regulator of G-protein signaling loco-like [Lucilia cuprina]
Woma_00003923-RA	XP_023297866.1 pickpocket protein 19-like [Lucilia cuprina]
Woma_00003988-RA	XP_023299481.1 dystrophin-like, partial [Lucilia cuprina]
Woma_00003822-RA	XP_023308456.1 attacin-A-like [Lucilia cuprina]
Woma_00004006-RA	XP_023298646.1 probable serine/threonine-protein kinase fhkB [Lucilia cuprina]
Woma_00004031-RA	XP_023304668.1 cGMP-dependent protein kinase, isozyme 2 forms cD4/T1/T3A/T3B-like [Lucilia cuprina]
Woma_00004203-RA	KNC25775.1 Protein croquemort, partial [Lucilia cuprina]
Woma_00004247-RA	XP_023301710.1 acyl-CoA synthetase family member 2, mitochondrial-like [Lucilia cuprina]
Woma_00004218-RA	XP_023307604.1 mucin-21-like [Lucilia cuprina]
Woma_00004289-RA	XP_012553862.1 PREDICTED: H(+)/Cl(-) exchange transporter 5-like [Hydra vulgaris]
Woma_00004280-RA	KNC25650.1 Serine protease easter [Lucilia cuprina]
Woma_00004277-RA	XP_023306909.1 lectin subunit alpha-like [Lucilia cuprina]
Woma_00004401-RA	XP_023292777.1 flocculation protein FLO11 [Lucilia cuprina]
Woma_00004404-RA	XP_023292773.1 neprilysin-1-like [Lucilia cuprina]
Woma_00004271-RA	XP_023301584.1 UDP-glucuronosyltransferase 2B15-like [Lucilia cuprina]
Woma_00004273-RA	XP_023301584.1 UDP-glucuronosyltransferase 2B15-like [Lucilia cuprina]
Woma_00004456-RB	XP_023297425.1 putative fatty acyl-CoA reductase CG5065 [Lucilia cuprina]
Woma_00004512-RA	XP_023298177.1 KDEL motif-containing protein 1-like [Lucilia cuprina]
Woma_00004521-RB	XP_023306188.1 putative mediator of RNA polymerase II transcription subunit 26 [Lucilia cuprina]
Woma_00004457-RA	XP_023297425.1 putative fatty acyl-CoA reductase CG5065 [Lucilia cuprina]
Woma_00004695-RA	XP_005176250.1 PREDICTED: PIH1 domain-containing protein 1 [Musca domestica]
Woma_00004742-RA	XP_014359147.1 PREDICTED: probable RNA-directed DNA polymerase from transposon BS [Papilio machaon]
Woma_00004702-RA	XP_023296311.1 zinc finger protein 391-like [Lucilia cuprina]
Woma_00004704-RA	XP_013102376.1 PREDICTED: zinc finger protein 782-like [Stomoxys calcitrans]
Woma_00006019-RA	XP_023303236.1 zinc finger protein 383-like [Lucilia cuprina]
Woma_00004839-RA	XP_023306955.1 collagen alpha-1(XI) chain-like [Lucilia cuprina]
Woma_00004840-RA	XP_023306955.1 collagen alpha-1(XI) chain-like [Lucilia cuprina]
Woma_00004803-RA	AMA07819.1 Jonah preproenzyme [Lucilia sericata]
Woma_00004804-RA	XP_023293222.1 collagenase-like [Lucilia cuprina]
Woma_00004949-RA	XP_023308585.1 rho guanine nucleotide exchange factor 25-like [Lucilia cuprina]
Woma_00004980-RA	XP_023306849.1 neuronal acetylcholine receptor subunit alpha-5-like [Lucilia cuprina]

6.2 Article 2

Jia, Z., Hasi, S., Zhan, D., Vogl, C., & Burger, P. A. (2024). Transcriptomic profiling of different developmental stages reveals parasitic strategies of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *BMC Genomics*, 25, 111. https://doi.org/10.1186/s12864-023-09949-3

Impact factor: 4.4

The data created within the article has been deposited at the following locations:

All raw transcriptome data in the study have been deposited in the National Center for Biotechnology Information's Sequence Read Archive (NCBI's SRA) database (https://www.ncbi.nlm.nih.gov/sra), including 4 second-stage larvae with accession numbers: SRX21712892, SRX21712893, SRX21712900, and SRX21712901, 4 third-stage larvae with accession numbers: SRX21712902, SRX21712903, SRX21712904, and SRX21712905, 4 pupae with accession numbers: SRX21712894, SRX21712895, SRX21712906, and SRX21712907, and 4 adult flies with accession numbers: SRX19591857, SRX19591858, SRX19591859, and SRX21712897. Supplementary materials, including Figure S1, Figure S2, Table S2, and Table S4, can be found immediately following the main text of this article in this thesis. Supplementary materials File S1, Table S1, Table S3, Table S5, Table S6, and Table S7 were deposited at Phaidra Vetmeduni with the identifier https://phaidra.vetmeduni.ac.at/o:2417.

BMC Genomics

RESEARCH Open Access

Transcriptomic profiling of different developmental stages reveals parasitic strategies of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly

Zhipeng Jia[1], Surong Hasi[2], Deng Zhan[2], Claus Vogl[3] and Pamela A. Burger[1*]

Abstract

Background *Wohlfahrtia magnifica* is an obligatory parasite that causes myiasis in several warm-blooded vertebrates. Adult females deposit the first-stage larvae directly onto wounds or natural body orifices (e.g., genitalia) of the host, from where they quickly colonize the host tissue and feed on it for development. The infestation of *W. magnifica* can lead to health issues, welfare concerns, and substantial economic losses. To date, little is known about the molecular mechanisms of the *W. magnifica*-causing myiasis.

Results In this study, we collected parasitic-stage larvae of *W. magnifica* from wounds of naturally infested Bactrian camels, as well as pupae and adult flies reared in vitro from the wound-collected larvae, for investigating the gene expression profiles of the different developmental stages of *W. magnifica*, with a particular focus on examining gene families closely related to the parasitism of the wound-collected larvae. As key proteins related to the parasite-host interaction, 2049 excretory/secretory (ES) proteins were identified in *W. magnifica* through the integration of multiple bioinformatics approaches. Functional analysis indicates that these ES proteins are primarily involved in cuticle development, peptidase activity, immune response, and metabolic processes. The global investigation of gene expression at different developmental stages using pairwise comparisons and weighted correlation network analysis (WGCNA) showed that the upregulated genes during second-stage larvae were related to cuticle development, peptidase activity, and RNA transcription and translation; during third-stage larvae to peptidase inhibitor activity and nutrient reservoir activity; during pupae to cell and tissue morphogenesis and cell and tissue development; and during adult flies to signal perception, many of them involved in light perception, and adult behavior, e.g., feeding, mating, and locomotion. Specifically, the expression level analysis of the likely parasitism-related genes in parasitic wound-collected larvae revealed a significant upregulation of 88 peptidase genes (including 47 serine peptidase genes), 110 cuticle protein genes, and 21 heat shock protein (hsp) genes. Interestingly, the expression of 2 antimicrobial peptide (AMP) genes, including 1 defensin and 1 diptericin, was also upregulated in the parasitic larvae.

Conclusions We identified ES proteins in *W. magnifica* and investigated their functional distribution. In addition, gene expression profiles at different developmental stages of *W. magnifica* were examined. Specifically, we focused

*Correspondence:
Pamela A. Burger
Pamela.Burger@vetmeduni.ac.at

Full list of author information is available at the end of the article

on gene families closely related to parasitism of wound-collected larvae. These findings shed light on the molecular mechanisms underlying the life cycle of the myiasis-causing fly, especially during the parasitic larval stages, and provide guidance for the development of control measures against *W. magnifica*.

Keywords Myiasis, *Wohlfahrtia magnifica*, Parasitic strategy, Excretory/Secretory (ES) proteins, RNA-seq, WGCNA

Introduction

Myiasis is the infestation of live vertebrates (humans or animals) with dipterous larvae [1]. During the parasitic larval stage, the larvae grow and develop by feeding on the host's tissue. The feeding activity damages the host tissue, resulting in significant morbidity, reduced animal welfare, and diminished milk and meat production and fertility [2]. Myiasis is a widespread problem, and flies causing myiasis are some of the world's most devastating insect pests, especially in poor countries with limited economic resources and health provisions. One study revealed that myiasis is the fourth most common travel-related skin disease in humans in non-endemic regions [3]. Myiasis can be categorized with respect to (i) the location of the parasitic site, e.g., a cutaneous, ophthalmic, ENT (Ear-nose-throat), intestinal, or urogenital site, or (ii) the relationship between the parasite and its host, which can be either obligatory, facultative, or accidental [1, 4].

Wohlfahrtia magnifica (Schiner, 1862) (Diptera: Sarcophagidae) is an obligatory traumatic (wound) myiasis-causing fly that can infest a range of mammals across Eastern and Southern Europe, Northern Africa, and Western and Northeast Asia [5–15]. For example, Remesar et al. examined 73,683 sheep from 122 flocks in the Albacete Province, Spain. They found a high overall prevalence of traumatic myiasis, with 95.9% at the flock level and 7.1% at the individual level [16]. Liu et al. surveyed 2038 female Bactrian camels in Inner Mongolia, China, between May and October 2021, and revealed an overall prevalence of 26.6% [17]. As a larviparous species, the female *W. magnifica* retains fertilized eggs inside her body. Once the first-stage larvae hatch, they are expelled onto or into wounds and natural body orifices (e.g., genitalia) of their host, where they start to develop by feeding on the host's tissue. As with other obligatory traumatic myiasis-causing flies, the parasitic activities of *W. magnifica* from the first- to the third-stage larvae can lead to health issues, animal welfare concerns, and significant economic losses [6, 18, 19]. In the subsequent non-parasitic post-larval stages, fully grown third-stage larvae leave the damaged tissue, fall to the ground, and burrow into the soil to pupate. Following a period of 10–15 days, adult flies emerge from pupae. About four days after mating, adult females actively search for a new host to lay the next generation of larvae, thus starting a new life cycle. During winter, in response to the cool weather, pupae enter a state of diapause, which is a period of dormancy.

During the parasitic stages, the myiasis-causing larvae employ sophisticated survival strategies. They penetrate the host's skin to invade the tissue, migrate within tissues, and feed on it. Some studies on myiasis-causing flies have demonstrated that proteolytic enzymes are a major component for the establishment, migration, feeding, growth, and development of myiasis-causing larvae. For instance, in *Lucilia cuprina*, in-vitro feeding assays showed that trypsin (a type of proteases) inhibitors led to significant larval growth retardation of first-stage larvae [20]. The activity study of *L. cuprina* collagenases and proteases suggested that they were not only involved in larval nutrition but also digested structural components of the skin, contributing to the formation of lesions and the production of exudates [21]. Similarly, Sandeman et al. also suggested that the proteases of *L. cuprina* degrade the host's tissue to facilitate migration and nutrition [22]. In *Oestrus ovis*, six major serine proteases secreted in the digestive tube of larvae were involved in larval trophic activity [23]. Furthermore, myiasis can give rise to necrosis, decay, and the excretion of host tissues, creating an environment conducive to the proliferation of bacteria [24]. Concurrently, the host's immune system is mobilized to mount an immune response to the invasion of the larvae. Thus, when parasitic larvae inhabit a wound, they actively defend against bacterial infection in the wounds, e.g., secreting antimicrobial peptides (AMPs), and modulate the host's immune response, e.g., natural killer (NK) cells and the complement system in the non-specific immune response, or antibodies and lymphocytes in the specific immune response, enabling larvae to establish a favorable environment for their survival and development. For example, as an active immune evasion strategy, *L. cuprina* secretes a protein called blowfly larval immunosuppressive protein (BLIP), which inhibits lymphocyte proliferation in sheep by at least 70%, compared with that in the presence of mitogen alone [25]. In *Hypoderma* spp., the proteases, *hypodermin A* (*HA*) and *hypodermin B* (*HB*), are involved in coping with specific and non-specific host immune systems [26]. HA and HB induce cleavage of the α and β chains of complement component 3 (C3) in naive bovine sera [27]. HA is also directly involved in the inhibition of lymphocyte proliferation in animals affected by hypodermosis [28, 29] and the cleavage of bovine IgG molecules into Fab and Fc fragments, with a reduction in the biological activity of these components [30].

From invading to leaving the host, larval gene expression patterns are attuned to the parasitic lifestyle. Investigating the gene expression profiles, such as peptidase genes and immune-related genes, at different developmental stages of *W. magnifica*, will facilitate the understanding of the complex molecular parasitism mechanisms of *W. magnifica* larvae. Transcriptome sequencing (such as RNA-seq) is an accurate, efficient, and economical method for studying gene expression at the genome-wide level. In combination with the genome of *W. magnifica*, which we recently sequenced and annotated [31], RNA-seq technology provides an excellent opportunity for analyzing the expression profiles of virtually all genes in *W. magnifica*. The ontogenetic transcriptional profiles can then be compared to those of numerous related insects exhibiting similar or different ecologies [32–34].

In this study, we identified the excretory/secretory (ES) protein collection in *W. magnifica*, which is involved in parasitism, and analyzed their functional distribution. Then, we sequenced the transcriptome of different developmental stages of *W. magnifica* using RNA-seq technology. Based on these RNA-seq data, we analyzed the global gene expression patterns at each developmental stage of *W. magnifica* using pairwise comparisons and weighted correlation network analysis (WGCNA) methods. Furthermore, with reference to the functional distribution of ES proteins, relevant literature related to myiasis-causing flies, and the nature of *W. magnifica* parasitism, we choose four specific parasitism-related gene families, including peptidase, cuticle protein, heat shock protein (hsp), and immune response genes, and looked into their expression patterns to gain a better understanding of how the larvae regulate expression of these specific genes to parasitize their hosts. Our findings provide valuable insights for the development of control strategies against *W. magnifica* infestations. This includes guiding the selection of potential targets for vaccines or insecticides aimed at disrupting the establishment of larvae on or in hosts.

Methods

Prediction of excretory/secretory (ES) proteins of *W. magnifica*

We identified the ES proteins from all protein sequences in *W. magnifica* using a workflow that integrates several tools. SignalP v. 5.0 [35] was employed to identify the classically secretory proteins, with the parameter for eukaryote organisms. The non-classical secretory proteins were predicted using SecretomeP v. 1.0 [36] with default options for mammalian organisms and further filtered by NN-scores larger than 0.9. The identified classical and non-classical secretory proteins were then subjected to a TMHMM v. 2.0 [37] with default parameters to detect transmembrane (TM) helices. The TM domain was further confirmed using the Phobius web server [38]. Among classical secretory proteins, those without a signal peptide and/or with hydrophobic helices of TM topologies that were distinguished from signal peptides were removed. Among non-classical secretory proteins, those with a signal peptide and/or TM helices were removed. Subsequently, mitochondrial proteins, proteins with the endoplasmic reticulum retention signal, and GPI-anchor proteins were identified using the TargetP v. 2.0 [39], the ScanProsite (Prosite pattern: PS00014) [40], and the PredGPI [41] web servers, respectively, and subsequently removed. The remaining proteins were considered to be predicted ES proteins in *W. magnifica*.

Functional annotation of ES proteins of *W. magnifica*

The eggNOG-mapper [42] was employed to obtain functional annotation of ES proteins. In terms of Gene Ontology (GO), the ES proteins were categorized into three high-level categories (Biological Processes, Cellular Component, and Molecular Function). The GO enrichment analysis was performed with the entire proteome as the reference group using padj <0.05 as the threshold for significant enrichment. Additionally, the KAAS v. 2.0 program [43] was applied to map the ES proteins to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, with the bi-directional best-hit (BBH) method for assigning orthologs.

Sample collection of *W. magnifica*

The subject of the study, *W. magnifica*, represents a non-endangered or protected invertebrate insect pest that affects agriculture. Second- and third-stage larvae of *W. magnifica* were non-invasively collected from the wounds of naturally infested domestic Bactrian camels in the field in Siziwang Banner, Ulanqab City, Inner Mongolia, China, with the consent and assistance of the owner. Therefore, no animal experimental or ethical permits were necessary. The experimental protocols adhered to the guidelines established by Inner Mongolia Agricultural University. The collected second-stage larvae were immediately dropped into liquid nitrogen to protect the RNA from degradation. The collected third-stage samples were divided into two groups. One group was carefully placed in a foam container filled with local soil for pupal and adult fly rearing, while the other group was submerged in liquid nitrogen. Three days after pupation a part of the individuals were stored at -80 °C. After a span of 14 days, the remaining pupae emerged into adult flies and were immediately frozen at -80 °C. In total, 16 samples, with four replicates for each of second-stage larvae, third-stage larvae, pupae, and adult flies, were collected for RNA extraction, library preparation, and RNA sequencing.

RNA isolation and assessment

The total RNA from all samples was extracted using the RNA Easy Fast Tissue/Cell kit (Tiangen Biotech, Beijing, China) following the manufacturer's instructions. To ensure compliance with the standards required for library preparation, the concentration, purity, and integrity of the extracted total RNA were measured using NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA), Agilent 5400 (Agilent Technologies, Palo Alto, CA, USA), and 1% agarose gels.

Illumina RNA-Seq library preparation, sequencing, and data filtering

After a quality check, Illumina RNA-seq libraries of 16 samples were prepared with the NEBNext® Ultra RNA Library Prep Kit for Illumina according to the manufacturer's protocol. The prepared libraries were then sequenced on an Illumina NovaSeq platform, generating 150 bp paired-end reads. Using BBduk in the BBTools toolset [44], the adaptor sequences, low-quality reads, and contamination from Bactrian camel and rRNA were removed and only reads of at least 75 bp in length were retained.

Analysis of gene expression of the different development stages

The resulting clean reads from 16 samples were used for the identification of differentially expressed genes (DEG) from pairwise comparisons between different developmental stages and WGCNA analysis. Clean data from three of these adult samples were also utilized in our separate article for the identification of DEGs between the sexes [45]. For each sample, the clean reads were aligned to the genome of *W. magnifica* [31] using the STAR program v. 2.2.1 [46], and read counts were quantified with the featureCounts program v.2.0.3 [47]. To improve the accuracy of gene expression, 5215 genes that had less than 30 reads across all samples were discarded. The raw counts of the remaining 11,506 genes served as input to the DESeq2 program v. 1.38.3 [48] in R for the analysis of DEG. A padj < 0.01 and an absolute log2FoldChange > 2 were set as cutoff criteria for the DEG assignment. The DEGs were subjected to the GO term enrichment analyses with a padj < 0.05 considered statistically significant.

Co-expression network analysis

A gene co-expression network was constructed using the WGCNA package v.1.72.1 [49] in R to assign the genes to the different modules. In order to reduce noise and improve module identification, genes with raw counts < 30 across all 16 samples were filtered. The remaining 11,506 genes were subject to the WGCNA program. Using the pickSoftThreshold function from the WGCNA package, a soft power of 22 was determined, resulting in a network with high mean connectivity and a coefficient of the scale free topology curve of $R^2 = 0.90$. The correlation network was created using the blockwiseModules function with the following parameters: power: 22, TOMType: signed, minModuleSize: 30, networkType: signed, mergeCutHeight: 0.35, and corType: pearson. The correlation of the trait (different development stages) to each module was calculated using the WGCNA function cor, and the respective significant values were generated using the WGCNA function corPvalueStudent. Gene significance (GS) and module membership (MM) were assessed for each module, where GS represented the association between gene expression profiles and each trait, and MM was defined as the correlation between gene expression profiles and each module eigengene (ME), with ME being the first principal component of the expression matrix of the corresponding module. The biological functions of genes in each of the modules were explored through GO enrichment analyses, applying a threshold of padj < 0.05. Furthermore, the hub genes with the top five intramodular connectivity degrees were selected from each module and analyzed their function.

Analysis of parasitism-associated gene families

ES proteins are closely associated with the successful parasitism of parasitic species. The study showed that the functions of the ES protein of *W. magnifica* are primarily distributed in cuticle development, peptidase activity, and immune response. Therefore, with reference to the functional distribution of ES protein in *W. magnifica*, along with literature on myiasis-causing flies, and the parasitic characteristics of *W. magnifica*, we selected four gene families, including peptidase, cuticle protein, heat shock protein (hsp), and immune response genes, and investigated their expression levels during the parasitic larval stage to understand how larvae regulate the expression of these specific genes for successful parasitization. To identify the members of the four families in the *W. magnifica* genome, the protein sequences of those families from *Drosophila melanogaster* were obtained from FlyBase (http://flybase.org, accessed on July 29, 2023) and used as query sequences to match the protein annotation file of the *W. magnifica* genome [31] using BLAST v. 2.14.0 [50] with an E-value threshold of < 1e−20.

Results

Prediction of ES proteins of *W. magnifica*

In this research, an array of bioinformatics tools was applied to predict the ES proteins of *W. magnifica*. Out of 16,718 putative proteins annotated in *W. magnifica*, a total of 2736 protein sequences were identified as potential ES protein candidates. Among them, 2458 were classified as classical ES proteins based on SignalP analysis, while 278 were characterized as non-classical ES

proteins as indicated by SecretomeP prediction. Using the TMHMM and Phobius programs, 614 proteins were excluded from this number: 559 classical secreted proteins lacked a signal peptide and/or contained at least one TM helix, while 55 proteins in the non-classical dataset had signal peptides and/or at least one TM helix (Fig. 1).

For the remaining 2122 putative ES proteins, the TargetP, the ScanProsite, and the PredGPI web servers were used to identify mitochondrial proteins, proteins with endoplasmic reticulum retention signal, and GPI-anchor proteins, respectively. Thus, one protein with the mitochondrial targeting signal, 21 proteins with the endoplasmic reticulum retention signal, and 51 proteins predicted to be GPI-anchored proteins were identified and removed. After this thorough screening, a total of 2049 proteins were recognized as ES proteins of *W. magnifica* (Fig. 1 and Supplementary File S1).

Functional annotation of ES proteins of *W. magnifica*

Among the 2049 ES proteins in *W. magnifica*, 1070 were assigned to a total of 6445 GO terms: 5179 from the Biological Process category, 504 from the Cellular Component category, and 762 from the Molecular Function category (Supplementary Table S1). Regarding the Biological Process category at the second level, the most prevalent GO terms included metabolic process, multicellular organismal process, cellular process, developmental process, biological regulation, response to stimulus, localization, immune system process, and reproduction (Fig. 2A). In the Cellular Component category, after removing redundancy, the second-level GO terms with the highest representation are extracellular region (Fig. 2A). At the second level of Molecular Function, binding, catalytic activity, structural molecule activity, and regulator activity terms are predominant, covering nearly 100% of all annotations (Fig. 2A). Within the third level of the two key second-level Molecular Functional terms, the top three terms under "binding" were protein binding, ion binding, and carbohydrate derivative binding (Fig. 2D), while under "catalytic activity", hydrolase activity and catalytic activity, acting on a protein terms are prevalent (Fig. 2E).

With respect to the KEGG pathway analysis, 382 out of 2049 ES proteins were found to be distributed across 246 pathways. The KEGG pathways with the highest representation are illustrated in (Fig. 2B), and the comprehensive dataset is available in (Supplementary Table S2). Among them, some of the pathways are likely associated with the parasitic life cycle of *W. magnifica* in hosts, such as immune response-related pathways (Toll and Imd signaling pathways) and digestion-related pathways (protein digestion and absorption and fat digestion and absorption).

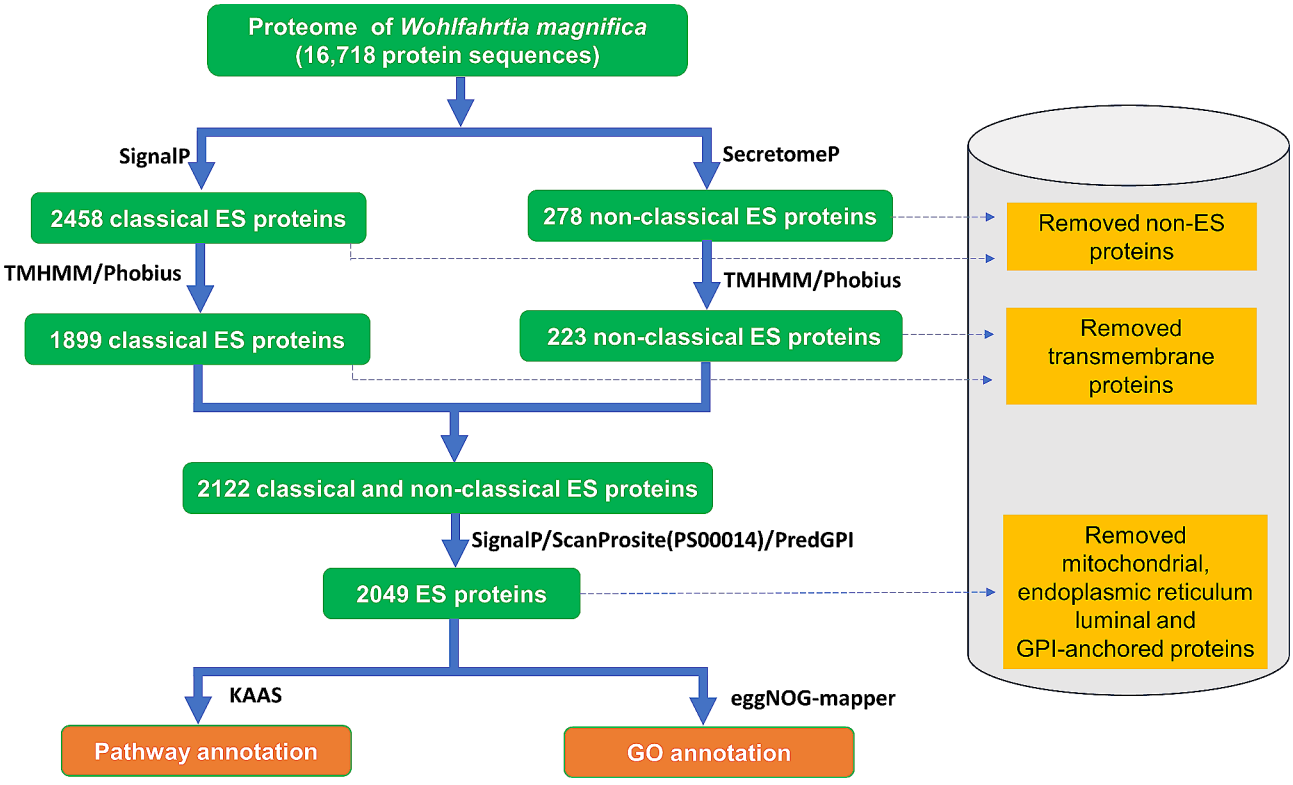

Fig. 1 Bioinformatic workflow used for ES protein analysis in *W. magnifica*

GO enrichment analysis of *W. magnifica* ES proteins

The GO term enrichment analysis provided valuable insights into the function distribution of ES proteins of *W. magnifica* (Supplementary Table S3). In the Biological Process category, the significant enrichment terms were related to proteolysis, developmental process (e.g., cuticle development), and immune response (e.g., response to bacterium) (Fig. 2C). In the Cellular Component category, the terms related to extracellular region showed significant enrichment (Fig. 2C). In the Molecular Function category, terms associated with catalytic activity (e.g., peptidase activity, peptidase inhibitor activity, and nutrient reservoir activity), binding (e.g., odorant binding), and structural constituent (e.g., structural constituent of cuticle), as well as receptor regulator activity displayed significant enrichment (Fig. 2C).

RNA-seq analysis

Summary statistics for the RNA-seq data

With the aim of analyzing DEGs across different developmental stages of *W. magnifica*, especially the parasitic larval stage, we sequenced second-stage larvae, third-stage larvae, pupae, and adult flies, with four replicates for each stage. Following the elimination of adaptor sequences, low-quality reads, and contamination stemming from Bactrian camel and rRNA, as well as the retention of only reads with a minimum length of 75 bp, the biological replicates of each developmental stage yielded a collection of clean reads spanning a range of 38 to 52 million 150-bp paired-end reads (Supplementary Table S4). Principal component analysis (PCA) revealed the presence of four distinct groupings corresponding to the different developmental stages (Supplementary Figure S1). In addition, we mapped the clean reads to the reference genome of *W. magnifica*, revealing that more than 95% of them could be uniquely or multiply mapped (Supplementary Table S4). These metrics indicated these clean reads with high quality. We noted that 1951 out of the 2049 ES protein genes had corresponding mapped mRNA-seq reads.

Analysis of differentially expressed genes (DEGs)

We discarded genes that had less than 30 reads across all developmental stage samples. The remaining 11,506 genes (including 1525 ES proteins) were analyzed for differential expression using the DEseq2 program in R. Our pairwise comparisons included: second-stage larvae versus third-stage larvae, second-stage larvae versus pupae, second-stage larvae versus adult flies, third-stage larvae versus pupae, third-stage larvae versus adult flies, and pupae versus adult flies.

Second-stage larvae versus third-stage larvae

When comparing second-stage larvae versus third-stage larvae, 574 genes were downregulated, while 523 genes were upregulated, which is with the smallest total number of DEGs in all pairwise comparisons (Fig. 3A). We performed an enrichment analysis of these DEGs to gain insight into the biological significance of these differences. GO terms dominantly enriched in downregulated DEGs are linked to structural constituent of cuticle in the Molecular Function category, and extracellular region in the Cellular Component category, as well as cuticle development in the Biological Process category. On the other hand, enrichment analysis of the upregulated DEGs revealed some of interesting GO terms associated with extracellular region and larval serum protein complex in the Cellular Components category, nutrient reservoir activity and peptidase inhibitor activity in the Molecular Function category, as well as response to bacterium, negative regulation of peptidase activity in the Biological Process category (Fig. 3A and Supplementary Table S5).

Second-stage larvae versus pupae and adult flies

We compared second-stage larvae to pupae and adult flies for DEG analysis. Between second-stage larvae and pupae, a total of 3201 genes were found to be significantly differentially expressed: 1781 genes were downregulated, and 1420 genes were upregulated (Fig. 3B). For the DEGs upregulated during second-stage larvae, a few top enriched GO terms are related to cuticle development. Other overrepresented terms were involved in small molecule metabolic process and transmembrane transporter activity. Also, we found interesting enriched terms connected to the parasitic life cycle in second-stage larvae, such as peptidase activity (e.g., endopeptidase activity, metallopeptidase activity, and exopeptidase activity). The DEGs upregulated during the pupal stage were enriched for cell and tissue development processes (Fig. 3B and Supplementary Table S5).

There are 2431 DEGs in the comparison between second-stage larvae and adult flies, consisting of 823 downregulated and 1608 upregulated genes (Fig. 3C). Consistent with previous comparisons, the majority of DEGs upregulated in second-stage larvae were enriched in cuticle development and peptidase activity. For DEGs upregulated in adult flies, a large portion of enriched GO terms belong to transmembrane transport and sensory perception, especially light perception. We also noted an enrichment of upregulated genes coding for behavior (e.g., reproductive behavior, feeding behavior, and locomotory behavior) (Fig. 3C and Supplementary Table S5).

Third-stage larvae versus pupae and adult flies

We also compared third-stage larvae to pupae and adult flies. Between third-stage larvae and pupae, a total of 2111 DEGs were identified, with 1196 genes downregulated and 915 genes upregulated (Fig. 3D). The GO enrichment analyses of the DEGs upregulated in

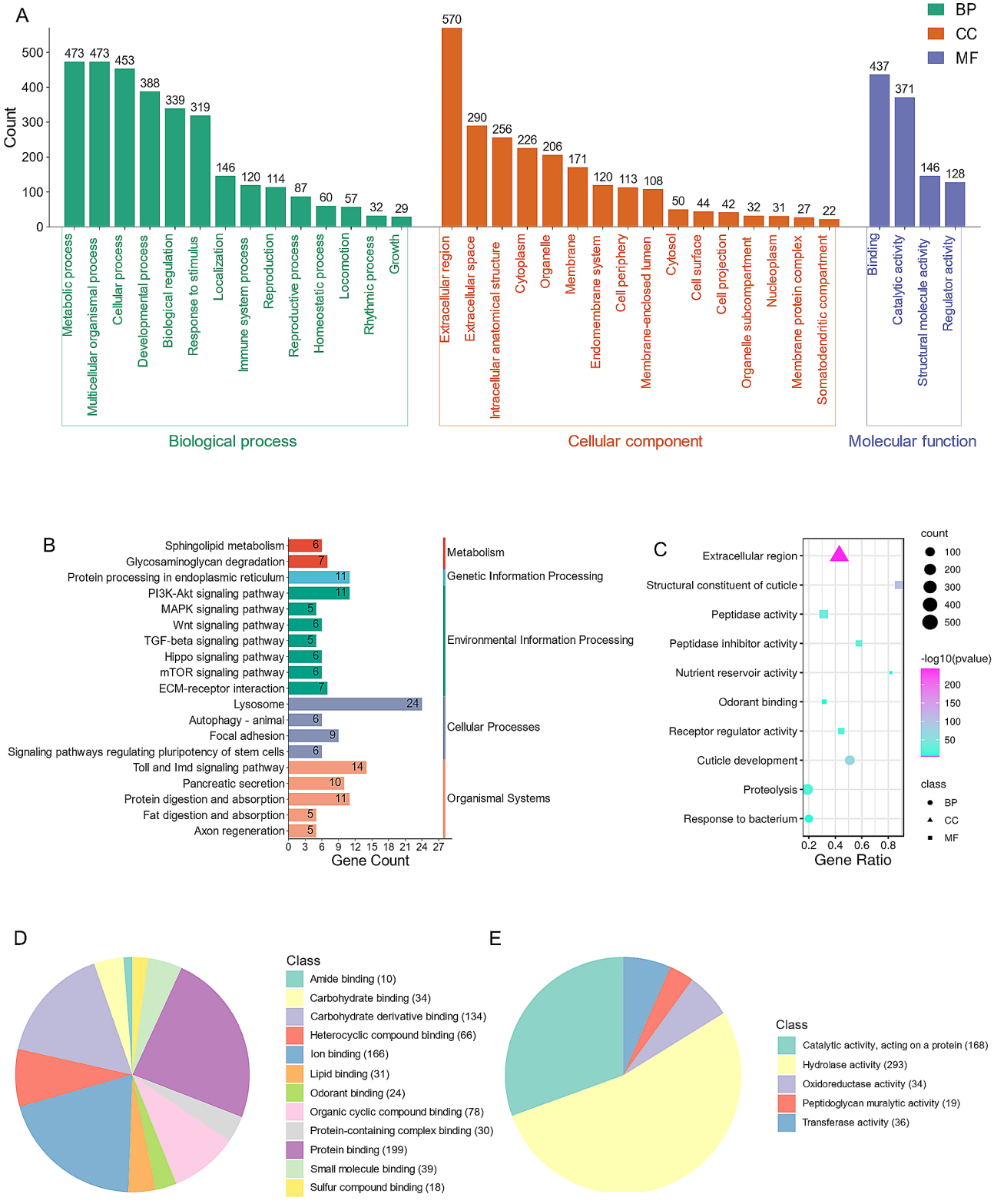

Fig. 2 Functional analysis of ES proteins of *W. magnifica*. (**A**) The distribution of GO terms at the second level. The horizontal axis represents the names of the GO terms, while the vertical axis represents the number of genes contained in each term. Green, orange, and blue colors represent the Biological Process, Cellular Component, and Molecular Function categories, respectively. (**B**) The distribution of KEGG pathways. The horizontal axis represents the number of genes contained in each KEGG pathway, while the vertical axis represents the names of the KEGG pathways. (**C**) The representative enriched GO terms. Triangles, squares, and circles indicate that the terms belong to the Cellular Component (CC), Molecular Function (MF), and Biological Process (BP) categories, respectively. (**D**) The terms in the third-level subcategory "binding" of Molecular Function. (**E**) The terms in the third-level subcategory "catalytic activity" of Molecular Function

third-stage larvae showed that some of GO terms were associated with cuticle development, peptidase activity, transmembrane transporter activity, response to nutrient, and small molecule catabolic process. Interestingly, we found that in third-stage larvae, the upregulated genes were also annotated for peptidase inhibitor activity, response to bacterium, and nutrient reservoir activity. The genes upregulated in the pupal stage were almost exclusively involved in the development of cells and tissues (Fig. 3D and Supplementary Table S5).

When comparing third-stage larvae and adult flies, a total of 2112 genes (690 downregulated and 1422 upregulated) exhibited significant changes in their expression levels (Fig. 3E). As in the comparison of third-stage larvae to pupae, a majority of the genes upregulated in third-stage larvae were enriched in cuticle development, peptidase activity, peptidase inhibitor activity, response to bacterium, cellular response to heat, and nutrient reservoir activity. On the other hand, DEGs related to transmembrane transporter activity, sensory perception (e.g., detection of light stimulus), and behavior were enriched in adult flies (Fig. 3E and Supplementary Table S5).

Pupae versus adult flies

Compared pupae and adult flies, we found that this transition is generally characterized by large-scale gene repression: the majority of DEGs (1856) were upregulated in the adult stage, while only 730 were upregulated in the pupal stage (Fig. 3F). GO term enrichment analysis was performed for DEGs between pupae and adult flies. A very large proportion of the upregulated DEGs in pupae were involved in the cell cycle process (e.g., Cellular Component: meiotic spindle and condensed chromosome; Molecular Function: microtubule binding, DNA-dependent ATPase activity, and ATP-dependent DNA helicase activity; Biological Process: mitotic nuclear division, mitotic spindle organization, and microtubule cytoskeleton organization involved in mitosis), and cells and tissues development (e.g., Molecular Function: sequence-specific DNA binding; Biological Process:

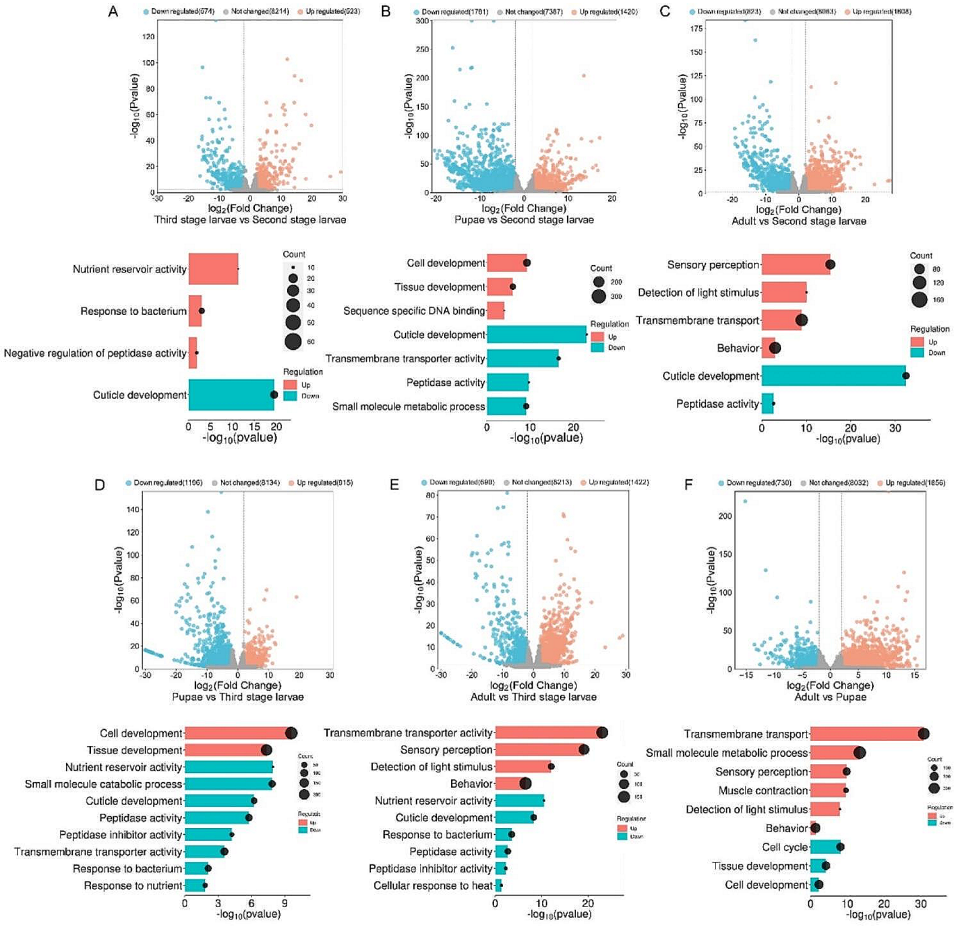

Fig. 3 Volcano plots (top) and the representative enriched GO terms (bottom) of upregulated and downregulated DEGs in pairwise comparisons across different developmental stages of *W. magnifica*. Red and blue colors represent upregulated and downregulated DEGs, respectively. (**A**) Second-stage larvae versus third-stage larvae. (**B**) Second-stage larvae versus pupae. (**C**) Second-stage larvae versus adult flies. (**D**) Third-stage larvae versus pupae. (**E**) Third-stage larvae versus adult flies. (**F**) Pupae versus adult flies

epithelium development, sensory organ development, eye development, wing disc development, nervous system development, and muscle organ development). Among the upregulated DEGs in adult flies, the top enriched GO terms were related to the transmembrane activity (e.g., Cellular Component: ion channel complex; Molecular Function: phosphate ion transmembrane transporter activity, amino acid transmembrane transporter activity, G-protein coupled receptor activity, channel activity, and symporter activity; Biological Process: ion transmembrane transport, organic acid transmembrane transport, amino acid transmembrane transport, and inorganic ion transmembrane transport) and sensory perception (e.g., Biological Process: detection of light stimulus, sensory perception of chemical stimulus, sensory perception of smell, and thermotaxis). We also observed some GO terms related to behavior (e.g., mating behavior and feeding behavior), muscle contraction, and small molecule metabolic process (Fig. 3F and Supplementary Table S5).

Overlapping analysis at different development stages

We analyzed the overlapping upregulated DEGs at each developmental stage. In comparisons of pupae versus second-stage larvae and adult flies versus second-stage larvae, 835 overlapping upregulated DEGs were identified. The GO enrichment analysis indicated that these genes were predominantly located in the extracellular region and functioned in peptidase activity and cuticle development (Fig. 4A and Supplementary Table S6). There were 500 overlapping upregulated DEGs between pupae versus third-stage larvae and adult flies versus third-stage larvae. The GO enrichment analysis showed that in addition to being enriched for similar functions as second-stage larvae in peptidase activity and cuticle development, DEGs were also enriched for functions in nutrient reservoir activity, defense response, and peptidase inhibitor activity (Fig. 4B and Supplementary Table S6).

There are 386 overlapping upregulated DEGs among second-stage larvae versus pupae, third-stage larvae versus pupae, and adult flies versus pupae. As expected, the functions of these overlapping genes in pupae primarily involve tissue morphogenesis, cell morphogenesis, and cell differentiation (Fig. 4C and Supplementary Table S6). In comparisons of second-stage larvae versus adult flies, third-stage larvae versus adult flies, and pupae versus adult flies, 756 upregulated DEGs were identified. These genes are functionally distributed in transmembrane transport activity, sensory perception, and behavior (Fig. 4D and Supplementary Table S6).

Interestingly, we found that the term "nutrient reservoir activity" was significantly enriched only in third-stage larvae (Fig. 4E).

Identification of modules specifically associated with different developmental stages

A total of 11,506 genes were allocated into twelve modules, each designated with a unique color to facilitate differentiation (Fig. 5A). Among them, a subset of twelve genes was assigned to the grey module, which included genes that could not be classified into any other specific module. The correlation between MEs and different developmental stages showed that five modules were significantly associated with specific traits, of which the brown module was correlated with second-stage larvae, the green and yellow modules with third-stage larvae, the turquoise module with pupae, and the blue module with adult flies (Fig. 5B). The GS and MM analysis demonstrated that the genes in each of the modules were correlated with the corresponding stages, confirming the fundamental importance of these modules in the network (Supplementary Figure S2).

Expanding our investigation, we performed the GO enrichment analysis for genes in each of these modules. In general, the results of the enrichment analysis of WGCNA were similar to those from the pairwise comparisons. However, there are some categories that are specifically detected by WGCNA. Similarly, the brown module related to second-stage larvae is primarily involved in cuticle development. We also identified a high number of categories associated with gene expression and protein synthesis (RNA processing, RNA binding, ribosome biogenesis, and translation), which were only observed to be enriched in WGCNA analysis (Fig. 5C and Supplementary Table S7). For third-larvae stage, the most enriched terms in the green and yellow modules were nutrient reservoir activity, response to bacterium, and small molecule catabolic process (Fig. 5D and Supplementary Table S7). The result of the pupae-correlated turquoise module is similar to the pairwise comparison, with numerous enriched terms implicated in the development of cells and tissues (Fig. 5E and Supplementary Table S7). The genes in the blue module related to adult flies were engaged in transmembrane transporter activity, sensory perception, and behaviors (Fig. 5F and Supplementary Table S7).

Hub genes within each module

We identified hub genes within each module. In the brown module associated with second-stage larvae, the top five genes in the core positions were Woma_00012859 (orthologs of TweedleG (*TwdlG*) of *D. melanogaster*), Woma_00004141 (DNA-directed RNA polymerase, mitochondrial (*Polrmt*)), Woma_00005556 (Acyl-CoA desaturase 1 (*Scd1*)), Woma_00007337 (3-hydroxy-3-methylglutaryl-coenzyme A reductase (*Hmgcr*)), and Woma_00003225 (Larval cuticle protein 65Ab1 (*Lcp65Ab1*)). Among them, Woma_00012859 and

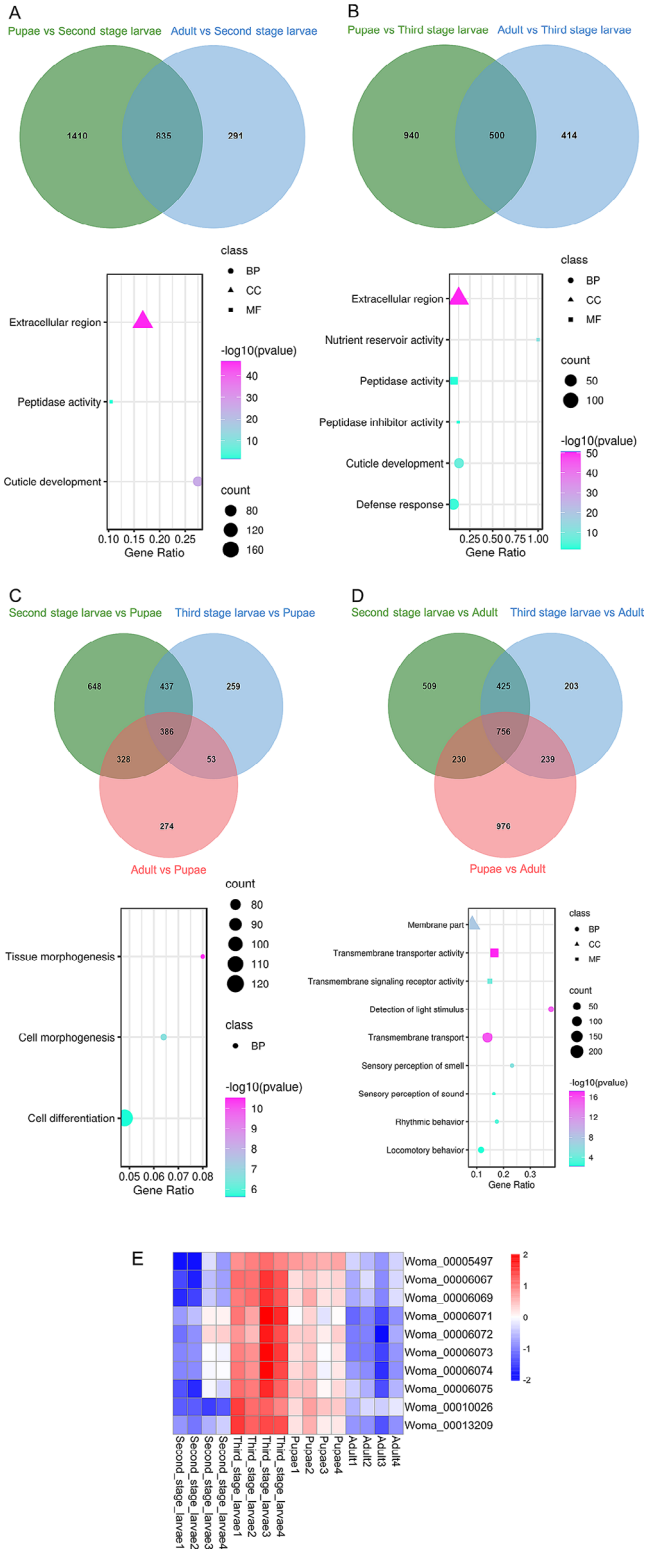

Fig. 4 The overlapping analysis across developmental stages of *W. magnifica*. (**A-D**) Venn diagram of overlapping DEGs (top) at each developmental stage and their representative enriched GO terms (bottom). Triangles, squares, and circles represent the terms belonging to the Cellular Component (CC), Molecular Function (MF), and Biological Process (BP) categories, respectively. (**A**) Second-stage larvae. (**B**) Third-stage larvae. (**C**) Pupae. (**D**) Adult flies. (**E**) Expression level heatmap of the genes in the GO term "nutrient reservoir activity" at each developmental stage. The horizontal axis represents the sample name; the vertical axis represents the gene names. Blue and red colors indicate low and high expression, respectively.

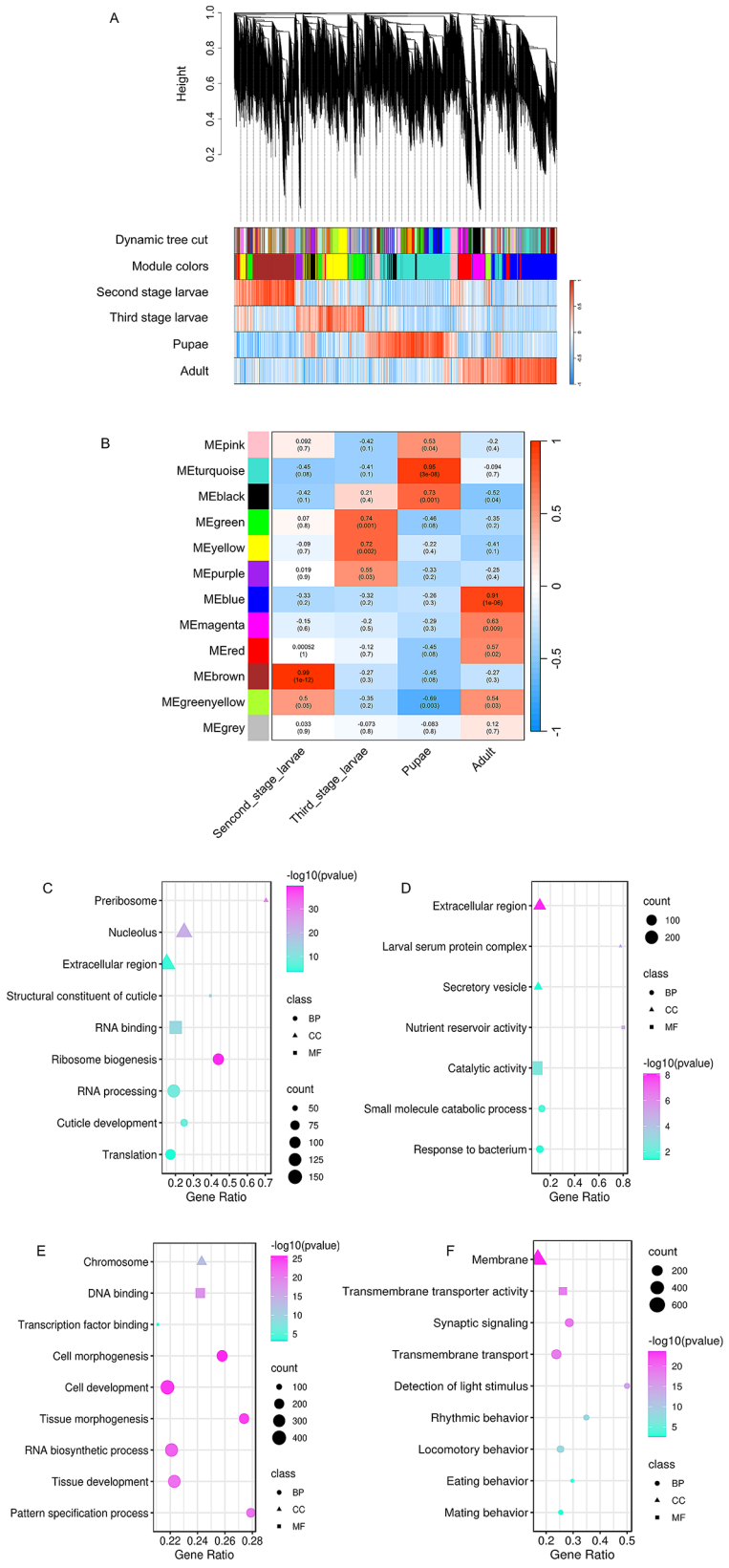

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 WGCNA analysis across developmental stages of *W. magnifica*. (**A**) Hierarchical clustering dendrogram of the genes. The expression distance is displayed on the vertical axis. The modules colored on the horizontal axis represent the branches of the clustering tree. The correlation of genes and modules, with red representing high correlation and blue representing low correlation. (**B**) The correlation of gene expression patterns for each module eigengene (ME) in relation to developmental stages. The rows correspond to developmental stages, and the columns represent MEs. Red represents a positive correlation, while blue indicates a negative correlation. (**C-F**) The representative enriched terms for genes in a module that exhibited a significant positive correlation with a specific developmental stage. Triangles, squares, and circles indicate that the terms belong to the Cellular Component (CC), Molecular Function (MF), and Biological Process (BP) categories, respectively. (**C**) The brown module (second-stage larvae). (**D**) The green and yellow modules (third-stage larvae). (**E**) The turquoise module (pupae). (**F**) The blue module (adult files)

Woma_00003225 are involved in cuticle development. The remaining three genes may be responsible for the growth and development of second-stage larvae.

There are two modules, green and yellow, related to third-stage larvae. The top five hub genes in the green module are Woma_00001427 (uncharacterized protein), Woma_00008050 (G-protein coupled receptor Mth2 (*mth2*)), Woma_00007519 (Aquaporin (*AQP*)), Woma_00005609 (Lysoplasmalogenase-like protein TMEM86A (*TMEM86A*)), and Woma_00013846 (uncharacterized protein). In the yellow module, the top five hub genes are Woma_00005806 (60 S ribosomal protein L37a (*RpL37A*)), Woma_00014656 (PRADC1-like protein), Woma_00015588 (Transmembrane emp24 domain-containing protein bai (*bai*)), Woma_00005411 (UPF0587 protein CG4646), and Woma_00009113 (larval cuticle protein 65Ab1-like). These proteins serve different functions. For example, Woma_00008050 is predicted to be involved in the G protein-coupled receptor signaling pathway and to function in response to starvation; Woma_00007519 regulates the permeability of cell membranes to water molecules.

In the turquoise module related to pupae, Woma_00006124 (Serine/arginine-rich splicing factor 7 (*SRSF7*)), Woma_00010385 (orthologs of CG18766 of *D. melanogaster*), Woma_00005980 (Protein neuralized (*neur*)), Woma_00009621 (Lamin-B receptor (*LBR*)), and Woma_00010498 (orthologs of CG6163 of *D. melanogaster*) were in the core positions. Some of them are involved in gene transcription. For example, Woma_00006124, an RS protein, plays a key role in precursor messenger RNA (pre-mRNA) splicing; Woma_00010385 is predicted to be involved in the positive regulation of DNA-templated transcription; Woma_00010498 is predicted to be involved in the regulation of transcription by RNA polymerase II. Woma_00009621 is a protein present in the nuclear membrane that interacts with nucleoskeletal proteins (e.g., lamin proteins) in the nucleus to maintain the stability and morphology of the nuclear membrane. It also participates in functions of the nucleus, such as gene transcription, DNA replication, and repair. Woma_00005980 is involved in the Notch signaling pathway. This pathway is essential for coordinating cell differentiation, tissue patterning, and organ development.

In the blue module linked to adult flies, the top five hub genes are Woma_00011427 (orthologs of CUB and LDLa domain (*Culd*) of *D. melanogaster*), Woma_00012405 (Guanine nucleotide-binding protein subunit gamma-e (*GBGE*)), Woma_00010769 (Sodium- and chloride-dependent GABA transporter 1 (*SLC6A1*)), Woma_00009321 (Carcinine transporter (*CarT*)), and Woma_00002878 (orthologs of retinophilin (*rtp*) of *D. melanogaster*). These genes are related to the response to light. Woma_00011427 encodes a photoreceptor-cell enriched transmembrane protein, which is required for the endocytic trafficking of the products of neither inactivation nor afterpotential E (*ninaE*) (functioning in light detection and vision) and transient receptor potential-like (*trpl*) (functioning in the response to light in photoreceptors). Woma_00012405 is a subunit of the G protein family, which serves as modulators or transducers in various transmembrane signaling systems. Woma_00010769 is required for the uptake of Gamma-aminobutyric acid (GABA) and other small molecules, and it likely plays an important role in establishing the excitatory/inhibitory balance of the central nervous system (CNS). Woma_00009321 encodes a transporter that is involved in the photoreceptor histamine-carcinine cycle. Woma_00002878 is a protein associated with visual function and is primarily found in photoreceptor cells in the retina. It plays an important role in maintaining visual adaptation, light signaling, and retinal function.

Expression level analysis of parasitism-related genes

With reference to the functional distribution of ES proteins of *W. magnifica*, relevant literature on other myiasis-causing flies, and the understanding of *W. magnifica*'s parasitic nature, we selected four gene families (peptidases, cuticle proteins, heat shock proteins (hsp), and immune response) and conducted an investigation into their expression patterns across different developmental stages, with special attention given to the parasitic larval stages.

Peptidase genes

We identified a total of 480 peptidase genes from *W. magnifica*, including 13 aspartic (A), 68 cysteine (C), 182 metallo- (M), 205 serine (S), and 12 threonine (T) peptidase genes.

We further investigated the expression levels of the peptidase genes across different developmental stages of *W. magnifica* (Fig. 6A). The results showed that 88

peptidase genes, including 6 aspartic, 6 cysteine, 28 metallo-, 47 serine, and 1 threonine, were upregulated during the second and/or third larval stages. At the pupal stage, the expression of only a few peptidase genes, including 0 aspartic, 4 cysteine, 4 metallo-, 9 serine, and 2 threonine, was upregulated in comparison to other developmental stages. At the adult stage, we identified 0 aspartic, 7 cysteine, 24 metallo-, 18 serine, and 2 threonine peptidase genes with upregulated expressions. Specifically, a number of studies have shown that serine peptidases are involved in parasitic activities, such as nutrient acquisition. Interestingly, in *W. magnifica*, up to 47 out of 205 serine peptidase genes were highly expressed during the parasitic larval stages. In contrast, 9 serine peptidase genes were highly expressed during the pupal stage, and 18 during the adult stage. For the remaining peptidase genes, expression levels in any of the three stages are not significantly higher than those in the other two stages.

Cuticle protein genes

A total of 215 cuticle protein genes were identified in the genome of *W. magnifica*. Interestingly, 196 out of 215 cuticle protein genes can be found within the ES protein collection.

We examined the expression of the cuticle protein genes across different developmental stages, and the corresponding expression heatmap is shown in (Fig. 6B). The results revealed that up to 51.16% (110/215) of the cuticle protein genes exhibited higher expression levels during the parasitic larval stage (second and/or third larval stages) compared to the pupal and adult stages. In contrast, only 2.79% (6/215) of the genes showed higher expression in the pupal stage than in other stages, and 13.95% (30/215) in the adult stage. The expression levels of the remaining 69 cuticle protein genes did not show a significant increase at any stage compared to the others.

Heat shock protein (hsp) genes

We identified 10 hsp genes in the HSP60 family, including 9 Chaperonin Containing TCP-1 (CCT) genes, 14 hsp genes in the HSP70 family, including 2 atypical hsp70 genes, 3 hsp genes in the HSP90 family, 1 hsp gene in the HSP100 family, and 16 hsp genes in the small HSP family in the genome of *W. magnifica*. Furthermore, in terms of co-factor chaperonins, we found 1 hsp gene in the HSP10 family and 38 hsp genes in the HSP40 family.

Subsequently, we examined the expression patterns of the hsp genes across the larval stage (second and/or third larval stages), the pupal stage, and the adult stage of *W. magnifica* (Fig. 6C). In the HSP60 family, only one gene, *Hsp60A*, exhibited upregulated expression during the larval stage compared to other developmental stages. Among the 14 genes in the HSP70 family, 9 hsp70 genes showed upregulated expressions during the larval stage compared to other developmental stages. For example, the expression levels of the *Hsc70-2* gene during the second larval stage were 390 times higher than those during the pupal stage and 21 times higher than those during the adult stage. Furthermore, 1 hsp70 gene, *Hsc70-1*, displayed higher expression during the adult stage compared to other developmental stages. Among the 16 hsp genes in the small HSP family, 4 small hsp genes displayed higher expressions during the larval stage compared to other developmental stages, 2 small hsp genes during the pupal stage, and 2 small hsp genes during the adult stage. In the HSP90 family, the only gene exhibited higher expression at the larval stage compared to other developmental stages.

Moreover, as for co-factor chaperonin families, the only gene in the HSP10 family, *Hsp10*, showed higher expression during the larval stage compared to other developmental stages; 5 hsp40 genes in the HSP40 family showed higher expression during the larval stage than during other developmental stages, 2 hsp40 genes during the pupal stage, and 1 hsp40 gene during the adult stage. There was no clear indication that the remaining hsp genes exhibited significantly higher expression at any stage compared to the others.

Immune response genes

We identified the core genes in four immune response pathways in *W. magnifica*, including the immune deficiency (Imd) pathway, responsible for Gram-negative bacteria infection; the Toll pathway, involved in fungus and Gram-positive bacterium infection; the c-Jun N-terminal kinase (JNK) pathway; and the Janus Kinase/Signal Transducer and Activator of Transcription (JAK/STAT) pathway. In the Imd and JAK/STAT pathways, all 18/18 and 6/6 core genes were identified, respectively, but none of the developmental stages exhibited higher expression than the others. In the Toll pathway, we identified 11/12 core genes. 1 gene, *spatzle* (*spz*), showed higher expression during the larval stage than during other developmental stages, 4 genes, *weckle* (*wek*), *Neurotrophin* 1 (*NT1*), and 2 *Toll-7* genes, during the pupal stage, and 1 gene, *spatzle 5* (*spz5*), during the adult stage. The JNK pathway includes 8/8 core genes, of which 3 genes—*puckered* (*puc*), *eiger* (*egr*), and *wengen* (*wgn*)—exhibited higher expression during the larval stage (second or third larval stages) compared to other developmental stages (Fig. 6D). The core genes that remained in the four pathways did not exhibit significantly upregulated expression at any specific stage compared to the others.

Additionally, our investigation revealed the presence of 21 AMPs in *W. magnifica*. The analysis of expression levels revealed that 1 defensin and 1 diptericin showed higher expression levels during the second larval stage compared to other developmental stages (Fig. 6E and F).

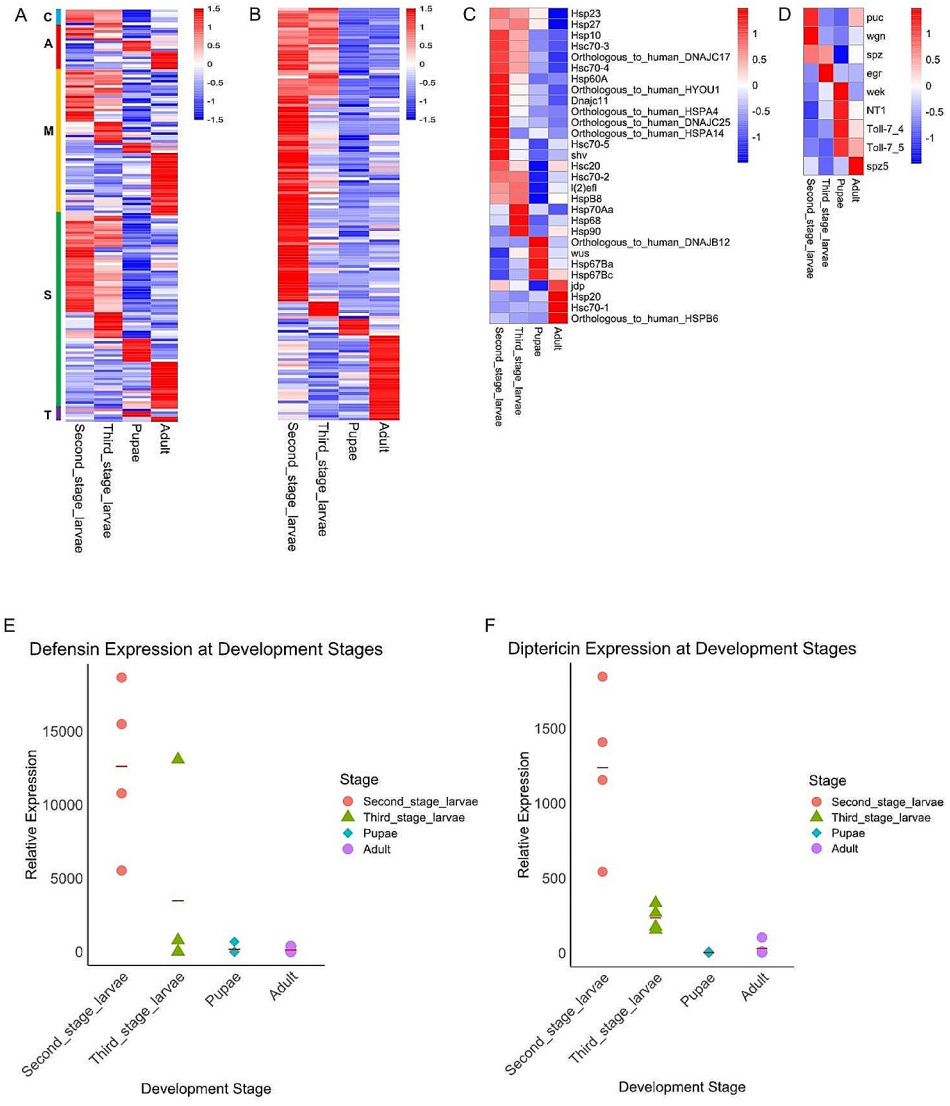

Fig. 6 Expression levels of genes associated with parasitism at different developmental stages of *W. magnifica*. (**A-D**) Expression level heatmap. The horizontal axis represents the sample names; the vertical axis represents the gene names. Blue and red colors indicate low and high expression levels, respectively. (**A**) Peptidase genes. C: Cysteine; A: Aspartic; M: Metallo-; S: Serine; T: Threonine. (**B**) Cuticle protein genes. (**C**) Hsp genes. (**D**) Immune response genes. (**E-F**) Expression level dot plot of antimicrobial peptide (AMP) genes. The horizontal axis represents the different developmental stages. The vertical axis represents the levels of gene expression. The horizontal red line represents the average value. (**E**) Defensin. (**F**). Diptericin

No significant changes in expression levels were detected in any one stage compared to the others among the remaining 19 AMPs.

Discussion

Excretory/secretory (ES) proteins are a class of proteins that some organisms release into their external environment. These proteins are of particular interest in the context of interactions between parasites and their hosts [51]. As a myiasis-causing insect pest, the identification of the *W. magnifica* ES proteins will facilitate the understanding of complex parasitic molecular mechanisms and contribute to the potential identification of drug or vaccine targets against the insect pest, as ES proteins within the extracellular environment may be more readily reachable by medications compared to other proteins. In our study, a collection of 2049 ES proteins from *W. magnifica* (accounting for 12.26% of the entire proteome) was identified using a pipeline that integrated a variety of bioinformatics approaches. The transcriptomic analysis in the study using RNA-seq data from different developmental stages showed that 95.22% of the ES protein genes have reads mapping, indicating that the ES protein genes are indeed expressed and suggesting the potential involvement of these ES proteins in the *W. magnifica* parasitic life cycle. We examined the functional distribution of

ES proteins. The GO functional annotation and enrichment analyses revealed that the top terms are associated with cuticle development, peptidase activity, immune response, and metabolic processes. Consistently, KEGG annotation assigned these genes to immune response and digestion-related pathways. These functions may be closely linked to successful parasitism of the larvae within the host.

We analyzed DEGs globally using both pairwise comparisons and WGCNA to more fully understand the life cycle of *W. magnifica* at the molecular level. A large number of cuticle protein and peptidase genes are highly expressed in the second larval stage. These genes are involved in larval molting and nutrient acquisition, respectively. Using WGCNA, we also identified a number of genes in relation to transcription and translation. In the life cycle of *W. magnifica*, growth and development of second-stage larvae are relatively rapid, which requires a large amount of protein synthesis. The gene expression and translation processes are key steps in protein synthesis, and therefore their expression may rise to respond to the cell's demand for proteins. The nutritional reserves of parasitic larvae are critical for the subsequent utilization of the free-living pupal and adult stages. For example, in the case of *O. ovis*, low larval weight (less than 280 mg) can affect the viability of pupal and adult stages [52]. At third-stage larvae of *W. magnifica*, genes related to nutrient reservoir activity and peptidase inhibitor activity were highly expressed. Therefore, the high expression of these genes indicates that nutrients accumulated in third-stage larvae for subsequent use by pupal and adult stages. The pupal stage is an essential stage in the transition from larvae to adult flies. As expected, during the pupal stage of *W. magnifica*, the pairwise comparisons and WGCNA results are consistent with a large number of genes associated with cell and tissue morphogenesis and cell and tissue development. In the adult stage, a large number of genes are involved in light perception. In insects, sensory signals are processed in the brain by dedicated neuronal circuits to guide behavior, such as light perception [53], which in particular has been examined in detail in *D. melanogaster* [54, 55]. The high expression of these genes in adult flies may be involved in adapting to the environment, finding food, mating, and regulating the circadian rhythm. In addition, there is a high expression of genes closely related to adult behaviors, such as eating behavior, mating behavior, and locomotory behavior. These behaviors are essential for the survival, reproduction, and adaptation of adult flies. In the study, while we successfully identified DEGs in second-stage larvae, third-stage larvae, pupae, and adult flies of *W. magnifica*, the acquisition of other earlier stages, such as eggs, from the field is challenging because currently cannot be reared successfully in laboratory conditions. Our next step involves adjusting the rearing environment and the feeding composition to establish a laboratory-rearing protocol for *W. magnifica*. This development would enable us to explore the gene expression patterns of other early developmental stages. Moreover, this would allow a comparative analysis of the DEGs between larvae from wounds of hosts and those from laboratory rearing, which would provide a clearer understanding of the specific genes involved in wound parasitism.

To further understand the parasitic mechanisms of *W. magnifica*, with reference to the functional distribution of ES proteins, relevant literature on myiasis-causing flies, and the parasitic characteristics of *W. magnifica*, we choose four likely parasitism-related gene families, including peptidase, cuticle protein, heat shock protein (hsp), and immune response genes, and investigated their expression patterns, especially the larvae collected from the host's wounds. Peptidases play an important role in the external protein digestion of parasitic myiasis-causing fly larvae, such as the establishment of infestation, wound formation, and nutrient acquisition in *L. cuprina* and *O. ovis*. In the study, we found that 88 out of 480 peptidase genes were highly expressed in the larval stage compared to other developmental stages of *W. magnifica*, while 19 and 51 out of 480 were highly expressed in the pupal and adult stages, respectively. In particular, for serine peptidases that are actively involved in parasitism in myiasis-causing flies [20–23, 56], up to 22.93% (47/205) were highly expressed in the parasitic larval stage, in comparison, only 4.39% (9/205) in the pupal stage and 8.78% (18/205) in the adult stage. These results suggest that a higher proportion of upregulated peptidase gene expression in the larval stage may contribute to tissue invasion and nutrient acquisition from hosts for the growth and development of larvae. Cuticles have a variety of key functions in the biology of insects, not only structuring their tough exoskeletons [57], but also serving as a barrier between the living tissues and the environment to protect them from dehydration, mechanical injury, predation, and insecticides [58, 59]. In *W. magnifica*, up to 51.16% (110/215) of cuticle protein genes are highly expressed during the larval stage, especially during the second larval stage. The large proportion of cuticle protein genes upregulated during larval stages shows the importance of cuticle proteins in growth and molting and may reflect adaptation to specific environments within the host's wounds. Insects defend themselves against viruses and bacteria using an innate immune system including cellular and humoral systems [60]. We investigated the immune pathway genes and antimicrobial peptides (AMPs) in *W. magnifica*. Interestingly, 1 defensin and 1 diptericin AMP genes were more highly expressed during the second larval stage than during other developmental stages. In insects, defensins are primarily active

against Gram-positive bacteria [61], while diptericin is active against Gram-negative bacteria [62]. As larval parasitism in wounds can cause severe damage to the host's tissues, making the environment susceptible to bacterial growth, the upregulated expression of these two AMPs may be involved in protecting larvae from both Gram-positive and Gram-negative bacterial infections. Moreover, we also found several hsp genes, such as 1 *Hsp60A* and 9 *Hsp70* genes, exhibited high expression in the larval stages. The body temperature of camels ranges from 34–40℃ [63]. In the study, samples of the larvae were obtained from the vaginas of female Bactrian camels. It is possible that the body temperature within the camel's wounds is higher than the temperature at which pupae develop optimally, thus leading to higher expression of hsp genes in response. Generally, our results demonstrate that *W. magnifica* responds to the complex environment within host wounds by employing multiple strategies involving the differential regulation of many genes and pathways. In the future, we can further explore genes within these four gene families that show increased expression during the parasitic larval stage. This will allow us to gain a deeper insight into the molecular mechanisms underlying the parasitism of *W. magnifica*. For example, in the context of peptidases, we can investigate which peptidase is involved in nutrient acquisition and which peptidase is responsible for evading the host's immune system.

Conclusions

W. magnifica is an obligatory parasite of several warm-blooded vertebrates and causes health and animal welfare problems and substantial economic loss. In our study, we identified a collection of ES proteins in *W. magnifica* that are closely related to parasitism. Functional analysis indicated that these ES proteins are involved in cuticle development, peptidase activity, immune response, and metabolic activities. An exploration using pairwise comparison and WGCNA methods revealed that during the second larval stage, genes closely associated with processes including peptidase activity and cuticle development were conspicuously upregulated; after molting to the third larval stage, gene expression patterns seem to be geared towards nutrient storage for utilization by pupae and adult flies; the pupal stage predominantly featured genes involved in cell and tissue morphogenesis and development orchestrating the transition from the larva to the adult fly; forwarding to the adult stage, genes exhibited a distinct tendency for signal perception, many of them implicated in light reception, and the behavioral activities of adult flies, such as feeding, mating, and locomotion. Specifically, when analyzing the expression profiles of genes tied to parasitism, a significant increase in gene expression was observed during the parasitic larval stage, including 88 peptidase (47 of which were serine peptidase), 110 cuticle protein, 21 hsp, and 2 AMPs genes, which may be targeted to engineer vaccines or pharmaceuticals to control *W. magnifica* or myiasis. Our findings further the understanding of the parasitic mechanism of *W. magnifica* and provide valuable opportunities to engineer control strategies against *W. magnifica*.

Abbreviations

AMP	Antimicrobial peptide
HA	Hypodermin A
HB	Hypodermin B
ES	Excretory/Secretory
WGCNA	Weighted correlation network analysis
HSP	Heat shock protein
TM	Transmembrane
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
DEG	Differentially expressed genes
GS	Gene significance
MM	Module membership
Imd	Immune deficiency
JNK	c-Jun N-terminal kinase
JAK/STAT	Janus kinase/signal transducer and activator of transcription

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-023-09949-3.

Supplementary Material 1: Supplementary Figure S1. Principal component analysis (PCA) of RNA-seq data from different developmental stages of *W. magnifica*. Red, green, blue, and purple dots represent samples of second-stage larvae, third-stage larvae, pupae, and adult flies, respectively

Supplementary Material 2: Supplementary Figure S2. Correlation between module membership (MM) and gene significance (GS) of the genes within each module. GS is plotted on the y-axis, and MM is plotted on the x-axis. The brown, green and yellow, tortoise, and blue dots represent genes in each of the modules. A The brown module to second-stage larvae. B-C the green and yellow modules to third-stage larvae. D The turquoise module to pupae. E The blue module to adult flies

Supplementary Material 3: Supplementary File S1. List of genes for excretory/secretory (ES) proteins of *W. magnifica*

Supplementary Material 4: Supplementary Table S1. GO annotation results for excretory/secretory (ES) proteins in *W. magnifica*

Supplementary Material 5: Supplementary Table S2. Assignment of the KEGG pathway for excretory/secretory (ES) proteins of *W. magnifica*

Supplementary Material 6: Supplementary Table S3. GO enrichment analysis of excretory/secretory (ES) proteins of *W. magnifica*

Supplementary Material 7: Supplementary Table S4. Statistics of transcriptome data at different developmental stages

Supplementary Material 8: Supplementary Table S5. GO enrichment analysis of differentially expressed genes in pairwise comparisons of different developmental stages of *W. magnifica*

Supplementary Material 9: Supplementary Table S6. GO enrichment analysis of overlapping DEGs between/among pairwise comparisons at each developmental stage of *W. magnifica*

Supplementary Material 10: Supplementary Table S7. GO enrichment analysis for genes in a module with a significant positive correlation with a specific developmental stage of *W. magnifica*

Acknowledgements
We gratefully acknowledge Professor Surong Hasi's laboratory members and the Bactrian camel owner for their assistance in sample collection.

Author contributions
ZJ and PAB conceived and designed the project. SH and DZ collected the samples. ZJ performed the data analysis. ZJ wrote the first draft of the manuscript. PAB, CV, and SH revised the manuscript. PAB, CV, and SH supervised the project. All authors read and approved the final manuscript.

Funding
Open Access funding provided by University of Veterinary Medicine Vienna. This work is supported by the Chinese Scholarship Council (CSC, the number 201909150004), the National Natural Science Foundation of China (the number 32060815) and the Austrian Science Funds (FWF) project (the number P29623-B25).

Data availability
All raw transcriptome data in the study have been deposited in the National Center for Biotechnology Information's Sequence Read Archive (NCBI's SRA) database (https://www.ncbi.nlm.nih.gov/sra), including four second-stage larvae with accession numbers: SRX21712892, SRX21712893, SRX21712900, and SRX21712901, four third-stage larvae with accession numbers: SRX21712902, SRX21712903, SRX21712904, and SRX21712905, four pupae with accession numbers: SRX21712894, SRX21712895, SRX21712906, and SRX21712907, and four adult flies with accession numbers: SRX19591857, SRX19591858, SRX19591859, and SRX21712897.

Declarations

Ethics approval and consent to participate
It is not relevant for the study as the research species, *Wohlfahrtia magnifica*, is an invertebrate agricultural insect pest, which is not an endangered or protected species. Where applicable, the research was conducted in compliance with institutional, national, and international guidelines.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Research Institute of Wildlife Ecology, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Savoyenstrasse 1, Vienna 1160, Austria
[2]Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture and Rural Affairs, College of Veterinary Medicine, Inner Mongolia Agricultural University, Hohhot 010018, China
[3]Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Veterinaerplatz 1, Vienna 1210, Austria

Received: 19 October 2023 / Accepted: 28 December 2023
Published online: 25 January 2024

References

1. Francesconi F, Lupi O. Myiasis. Clin Microbiol Rev. 2012;25(1):79–105.
2. Zumpt F. Myiasis in man and animals in the Old World. London: Butterworths; 1965.
3. Caumes E, Carrière J, Guermonprez G, Bricaire F, Danis M, Gentilini M. Dermatoses associated with travel to tropical countries: a prospective study of the diagnosis and management of 269 patients presenting to a Tropical Disease unit. Clin Infect Dis. 1995;20(3):542–48.
4. Noutsis C, Millikan LE, Myiasis. Dermatol Clin. 1994;12(4):729–36.
5. Valentin A, Baumann MPO, Schein E, Bajanbileg S. Genital myiasis (wohlfahrtiosis) in camel herds of Mongolia. Vet Parasitol. 1997;73(3–4):335–46.
6. Farkas R, Hall MJR, Kelemen F. Wound myiasis of sheep in Hungary. Vet Parasitol. 1997;69(1–2):133–44.
7. Gaglio G, Brianti E, Abbene S, Giannetto S. Genital myiasis by *Wohlfahrtia Magnifica* (Diptera, Sarcophagidae) in Sicily (Italy). Parasitol Res. 2011;109:1471–74.
8. Giangaspero A, Traversa D, Trentini R, Scala A, Otranto D. Traumatic myiasis by *Wohlfahrtia Magnifica* in Italy. Vet Parasitol. 2011;175(1–2):109–12.
9. Hall MJR, Testa JM, Smith L, Adams ZJO, Khallaayoune K, Sotiraki S, et al. Molecular genetic analysis of populations of Wohlfahrt's wound myiasis fly, *Wohlfahrtia Magnifica*, in outbreak populations from Greece and Morocco. Med Vet Entomol. 2009;23:72–9.
10. Ruiz Martínez I, Leclercq M. Data on distribution of screwworm fly *Wohlfahrtia Magnifica* (Schiner) in southwestern Europe (Diptera: Sarcophagidae). Notes Fauniques Gembloux. 1994;28:53–60.
11. Sotiraki S, Farkas R, Hall MJR. Fleshflies in the flesh: epidemiology, population genetics and control of outbreaks of traumatic myiasis in the Mediterranean Basin. Vet Parasitol. 2010;174(1–2):12–8.
12. Yasuda M. Morphology of the Larva of *Wohlfahrtia Magnifica* Schin. Found in a wound on a Camel in Inner Mongolia. J Chosen Nat Hist Soc. 1940;7(29):27–36.
13. Farkas R, Hall MJR, Bouzagou AK, Lhor Y, Khallaayoune K. Traumatic myiasis in dogs caused by *Wohlfahrtia Magnifica* and its importance in the epidemiology of wohlfahrtiosis of livestock. Med Vet Entomol. 2009;23:80–5.
14. Diakakis N, Papadopoulos E, Hall MJR, Desiris A. Post-traumatic complication due to *Wohlfahrtia magnifica* larvae on a horse. Vet Rec. 2006;158(5):170–72.
15. Schumann H, Ribbeck R, Beulig W. *Wohlfahrtia Magnifica* (Schiner, 1862) Diptera: Sacrophagidae) causing a vaginal myiasis in domesticated two-humped camels in the Mongolian people's republic. Archiv fur Experimentelle Veterinarmedizin. 1976;30(6):799–806.
16. Remesar S, Otero JL, Panadero R, Díez-Baños P, Díaz P, García-Díos D, et al. Traumatic myiasis by *Wohlfahrtia Magnifica* in sheep flocks from southeastern Spain: prevalence and risk factors. Med Vet Entomol. 2022;36(1):30–7.
17. Liu J, Hou B, Wuen J, Jiang N, Gao T, Hasi S. Epidemiological investigation on genital myiasis of bactrian camels in parts of Inner Mongolia, China. J Camel Pract Res. 2022;29:229–35.
18. İpek DNS, Şaki CE, Çay M. The investigation of lipid peroxidation, anti-oxidant levels and some hematological parameters in sheep naturally infested with *Wohlfahrtia magnifica* larvae. Vet Parasitol. 2012;187(1–2):112–18.
19. Martinez RI, Cruz SMD, Rodriguez R, Lopez DM, Parra MS, Navio FA. Myiasis caused by *wohlfahrtia magnjfica* in southern Spain. Isr J Vet Med. 1987;43(1):34–41.
20. Casu RE, Jarmey JM, Elvin CM, Eisemann CH. Isolation of a trypsin-like serine protease gene family from the sheep blowfly *Lucilia Cuprina*. Insect Mol Biol. 1994;3(3):159–70.
21. Bowles VM, Carnegie PR, Sandeman RM. Characterization of proteolytic and collagenolytic enzymes from the larvae of *Lucilia Cuprina*, the sheep blowfly. Aust J Biol Sci. 1988;41(2):269–78.
22. Sandeman RM, Feehan JP, Chandler RA, Bowles VM. Tryptic and chymotryptic proteases released by larvae of the blowfly, *Lucilia Cuprina*. Int J Parasitol. 1990;20(8):1019–23.
23. Tabouret G, Bret-Bennis L, Dorchies P, Jacquiet P. Serine protease activity in excretory–secretory products of *Oestrus ovis* (Diptera: Oestridae) larvae. Vet Parasitol. 2003;114(4):305–14.
24. Erdmann GR. Antibacterial action of myiasis-causing flies. Parasitol Today. 1987;3(7):214–16.
25. Elkington RA, Humphries M, Commins M, Maugeri N, Tierney T, Mahony TJ. A *Lucilia Cuprina* excretory–secretory protein inhibits the early phase of lymphocyte activation and subsequent proliferation. Parasite Immunol. 2009;31(12):750–65.
26. Moiré N. Hypodermin A and inhibition of lymphocyte proliferation. Parasitol Today. 1998;14(11):455–57.
27. Boulard C. Degradation of bovine C3 by serine proteases from parasites *Hypoderma lineatum* (Diptera, Oestridae). Vet Immunol Immunopathol. 1989;20(4):387–98.
28. Chabaudie N, Boulard C. Effect of hypodermin A, an enzyme secreted by *Hypoderma lineatum* (insect Oestridae), on the bovine immune system. Vet Immunol Immunopathol. 1992;31(1–2):167–77.
29. Moiré N, Nicolas-Gaulard I, Le Vern Y, Boulard C. Enzymatic effect of hypodermin A, a parasite protease, on bovine lymphocyte membrane antigens. Parasite Immunol. 1997;19(1):21–7.

30. Pruett JH Jr. Proteolytic cleavage of bovine IgG by hypodermin A, a serine protease of *Hypoderma lineatum* (Diptera: Oestridae). J Parasitol. 1993;79(6):829–33.
31. Jia Z, Hasi S, Vogl C, Burger PA. Genomic insights into evolution and control of *Wohlfahrtia Magnifica*, a widely distributed myiasis-causing fly of warm-blooded vertebrates. Mol Ecol Resour. 2022;22(7):2744–57.
32. Scott MJ, Benoit JB, Davis RJ, Bailey ST, Varga V, Martinson EO, et al. Genomic analyses of a livestock pest, the New World screwworm, find potential targets for genetic control programs. Commun Biol. 2020;3(1):424.
33. Kim JY, Lim HY, Shin SE, Cha HK, Seo JH, Kim SK, et al. Comprehensive transcriptome analysis of *Sarcophaga Peregrina*, a forensically important fly species. Sci data. 2018;5(1):1–8.
34. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, et al. The developmental transcriptome of *Drosophila melanogaster*. Nature. 2011;471(7339):473–79.
35. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37(4):420–23.
36. Bendtsen JD, Jensen LJ, Blom N, von Heijne G, Brunak S. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel. 2004;17(4):349–56.
37. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567–80.
38. Käll L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids res. 2007;35(suppl2):W429–W32.
39. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 2000;300(4):1005–16.
40. De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids res. 2006;34(suppl2):W362–W65.
41. Pierleoni A, Martelli PL, Casadio R. PredGPI: a GPI-anchor predictor. BMC Bioinformatics. 2008;9(1):1–11.
42. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38(12):5825–29.
43. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res (suppl2). 2007;35:W182–W5.
44. Bushnell B, BBTools. A suite of fast, Multithreaded Bioinformatics Tools Designed for Analysis of DNA and RNA sequence data. Joint Genome Institute; 2018.
45. Jia Z, Hasi S, Zhan D, Hou B, Vogl C, Burger PA. Genome and Transcriptome Analyses Facilitate Genetic Control of *Wohlfahrtia Magnifica*, a myiasis-causing flesh fly. Insects. 2023;14(7):620.
46. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
47. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
48. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.
49. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):1–13.
50. Ye J, McGinnis S, Madden TL. BLAST: improvements for better sequence analysis. Nucleic Acids Res. 2006;34(suppl2):W6–W9.
51. Okakpu OK, Dillman AR. Review of the role of parasitic nematode excretory/secretory proteins in host immunomodulation. J Parasitol. 2022;108(2):199–208.
52. Cepeda-Palacios R, Frugère S, Dorchies P. Expected effects of reducing *Oestrus ovis* L. mature larval weight on adult populations. Vet Parasitol. 2000;90(3):239–46.
53. Behnia R, Desplan C. Visual circuits in flies: beginning to see the whole picture. Curr Opin Neurobiol. 2015;34:125–32.
54. Paulk A, Millard SS, van Swinderen B. Vision in Drosophila: seeing the world through a model's eyes. Annu Rev Entomol. 2013;58:313–32.
55. Zuker CS. The biology of vision of Drosophila. Proc Natl Acad Sci U S A. 1996;93(2):571–76.
56. Casu RE, Eisemann CH, Vuocolo T, Tellam RL. The major excretory/secretory protease from *Lucilia Cuprina* larvae is also a gut digestive protease. Int J Parasitol. 1996;26(6):623–28.
57. Andersen SO. Biochemistry of insect cuticle. Annu Rev Entomol. 1979;24(1):29–59.
58. Muthukrishnan S, Mun S, Noh MY, Geisbrecht ER, Arakane Y. Insect cuticular chitin contributes to form and function. Curr Pharm Des. 2020;26(29):3530–45.
59. Balabanidou V, Grigoraki L, Vontas J. Insect cuticle: a critical determinant of insecticide resistance. Curr Opin Insect Sci. 2018;27:68–74.
60. Strand MR. The insect cellular immune response. Insect sci. 2008;15(1):1–14.
61. Hoffmann JA, Hetru C. Insect defensins: inducible antibacterial peptides. Immunol Today. 1992;13(10):411–15.
62. Keppi E, Pugsley AP, Lambert J, Wicker C, Dimarcq JL, Hoffmann JA, et al. Mode of action of diptericin A, a bactericidal peptide induced in the hemolymph of *Phormia terranovae* larvae. Arch Insect Biochem Physiol. 1989;10(3):229–39.
63. Schmidt-Nielsen K, Schmidt-Nielsen B, Jarnum SA, Houpt TR. Body temperature of the camel and its relation to water economy. Am J Physiology-Legacy Content. 1956;188(1):103–12.

Publisher's Note

Supplementary Materials for:

Transcriptomic profiling of different developmental stages reveals parasitic strategies of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly

Zhipeng Jia[1], Surong Hasi[2], Deng Zhan[2], Claus Vogl[3], Pamela A. Burger[1*]

1. Research Institute of Wildlife Ecology, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Savoyenstrasse 1, 1160, Vienna, Austria
2. Inner Mongolia Agricultural University/Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture and Rural Affairs, Hohhot, 010018, China
3. Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria
*Corresponding author: pamela.burger@vetmeduni.ac.at

The following contents are included:

Supplementary Figure S1 and Figure S2
Supplementary Table S2 and Table S4

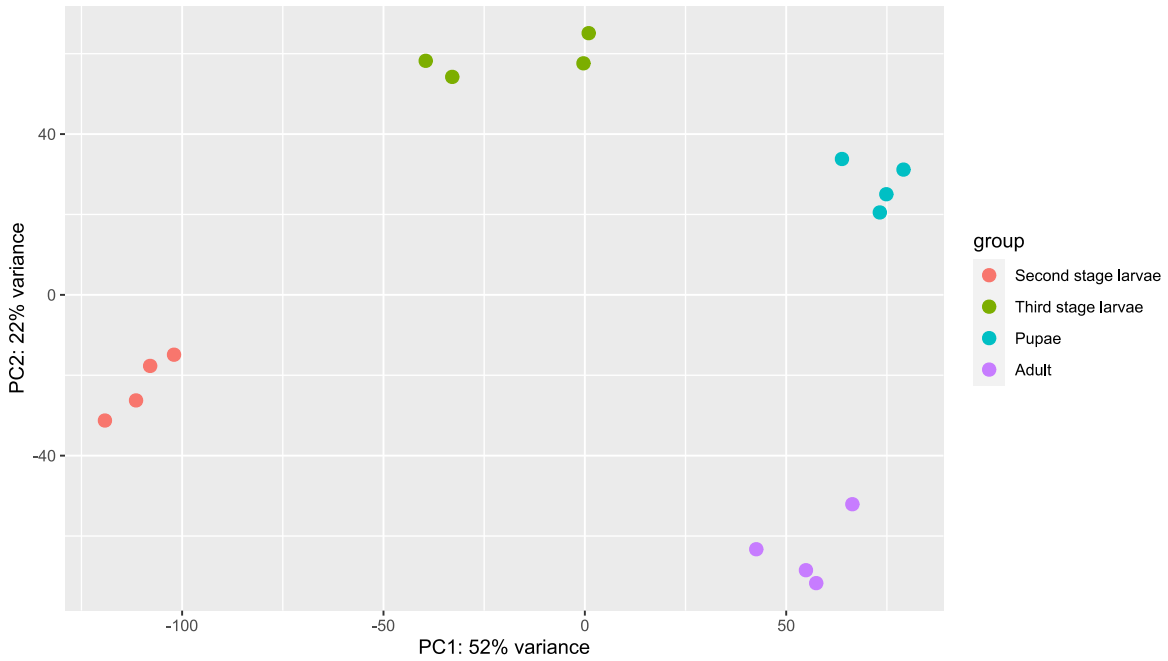

Figure S1. Principal component analysis (PCA) of RNA-seq data from different developmental stages of *W. magnifica*. Red, green, blue, and purple dots represent samples of second-stage larvae, third-stage larvae, pupae, and adult flies, respectively.

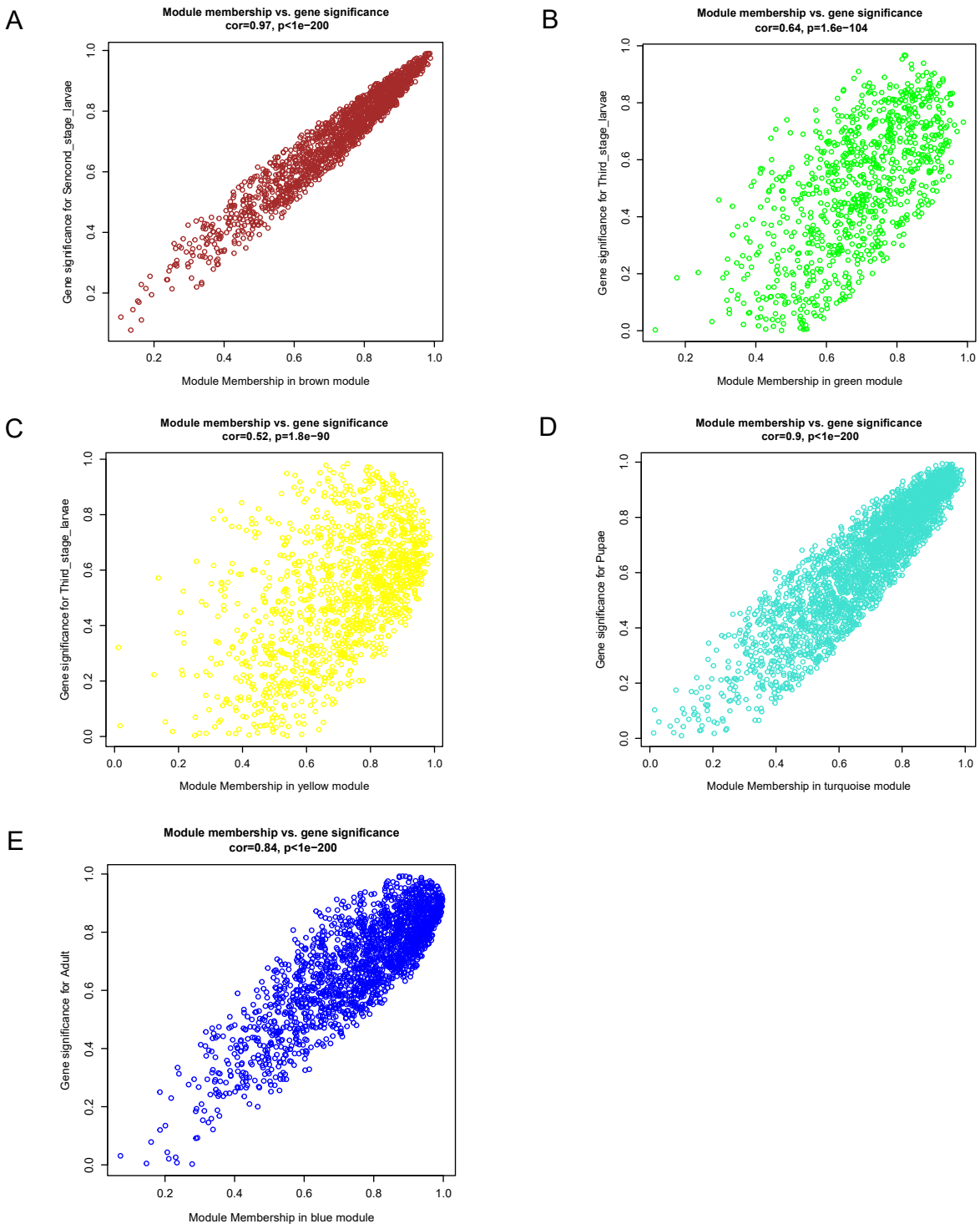

Figure S2. Correlation between module membership (MM) and gene significance (GS) of genes within each module. MM is plotted on the x-axis, and GS is plotted on the y-axis. The brown, green and yellow, tortoise, and blue dots represent genes in each of the modules. **A** The brown module to second-stage larvae. **B-C** The green and yellow modules to third-stage larvae. **D** The turquoise module to pupae. **E** The blue module to adult.

Table S2. Assignment of the KEGG pathway for secretory/excretory (ES) proteins of *W. magnifica*.

Pathway name	Gene counts
Lysosome	24
Pathways in cancer	16
Toll and Imd signaling pathway	14
Human papillomavirus infection	14
Protein processing in endoplasmic reticulum	11
PI3K-Akt signaling pathway	11
Protein digestion and absorption	11
Pancreatic secretion	10
Focal adhesion	9
Proteoglycans in cancer	8
Breast cancer	8
Pathways of neurodegeneration - multiple diseases	8
Glycosaminoglycan degradation	7
ECM-receptor interaction	7
Alzheimer disease	7
Diabetic cardiomyopathy	7
Sphingolipid metabolism	6
Other glycan degradation	6
Wnt signaling pathway	6
Hippo signaling pathway	6
mTOR signaling pathway	6
Autophagy - animal	6
Signaling pathways regulating pluripotency of stem cells	6
Hepatocellular carcinoma	6
Amyotrophic lateral sclerosis	6
Prion disease	6
Other types of O-glycan biosynthesis	5
MAPK signaling pathway	5
TGF-beta signaling pathway	5
Fat digestion and absorption	5
Axon regeneration	5
Longevity regulating pathway - multiple species	5
Gastric cancer	5
Basal cell carcinoma	5
Amoebiasis	5
Non-alcoholic fatty liver disease	5
Cushing syndrome	5
Purine metabolism	4
Drug metabolism - other enzymes	4
MAPK signaling pathway - fly	4
Ras signaling pathway	4
Rap1 signaling pathway	4
HIF-1 signaling pathway	4
Sphingolipid signaling pathway	4
cAMP signaling pathway	4
AMPK signaling pathway	4
Apoptosis	4
Adherens junction	4
Antigen processing and presentation	4
Melanogenesis	4
Chemical carcinogenesis - receptor activation	4
Chemical carcinogenesis - reactive oxygen species	4
Small cell lung cancer	4

Parkinson disease	4
Endocrine resistance	4
Galactose metabolism	3
Glycerolipid metabolism	3
Glycerophospholipid metabolism	3
Ubiquitin mediated proteolysis	3
Hippo signaling pathway - fly	3
Neuroactive ligand-receptor interaction	3
Phagosome	3
Regulation of actin cytoskeleton	3
Fc gamma R-mediated phagocytosis	3
Insulin signaling pathway	3
Thyroid hormone signaling pathway	3
Renin-angiotensin system	3
Cholesterol metabolism	3
Axon guidance	3
Longevity regulating pathway	3
Thermogenesis	3
Transcriptional misregulation in cancer	3
Colorectal cancer	3
Prostate cancer	3
Human immunodeficiency virus 1 infection	3
Influenza A	3
Human cytomegalovirus infection	3
Shigellosis	3
Toxoplasmosis	3
Huntington disease	3
Lipid and atherosclerosis	3
Fluid shear stress and atherosclerosis	3
Insulin resistance	3
Glycolysis/Gluconeogenesis	2
Pentose and glucuronate interconversions	2
Ascorbate and aldarate metabolism	2
Starch and sucrose metabolism	2
Amino sugar and nucleotide sugar metabolism	2
Pyruvate metabolism	2
Oxidative phosphorylation	2
Tyrosine metabolism	2
Glutathione metabolism	2
Glycosaminoglycan biosynthesis - chondroitin sulfate/dermatan sulfate	2
Glycosylphosphatidylinositol GPI-anchor biosynthesis	2
Glycosphingolipid biosynthesis - ganglio series	2
Porphyrin metabolism	2
Insect hormone biosynthesis	2
Polycomb repressive complex	2
Viral life cycle - HIV-1	2
Notch signaling pathway	2
Hedgehog signaling pathway - fly	2
Apelin signaling pathway	2
FoxO signaling pathway	2
Cytokine-cytokine receptor interaction	2
Endocytosis	2
Oocyte meiosis	2
Necroptosis	2
Motor proteins	2
Th1 and Th2 cell differentiation	2
Insulin secretion	2
Glucagon signaling pathway	2

Ovarian steroidogenesis	2
Progesterone-mediated oocyte maturation	2
Parathyroid hormone synthesis, secretion and action	2
Renin secretion	2
Cardiac muscle contraction	2
Salivary secretion	2
Bile secretion	2
Vitamin digestion and absorption	2
Dorso-ventral axis formation	2
Circadian entrainment	2
MicroRNAs in cancer	2
Viral carcinogenesis	2
Choline metabolism in cancer	2
Pancreatic cancer	2
Thyroid cancer	2
Acute myeloid leukemia	2
Renal cell carcinoma	2
Coronavirus disease - COVID-19	2
Kaposi sarcoma-associated herpesvirus infection	2
Epstein-Barr virus infection	2
Salmonella infection	2
Hypertrophic cardiomyopathy	2
Arrhythmogenic right ventricular cardiomyopathy	2
Viral myocarditis	2
Type II diabetes mellitus	2
Alcoholic liver disease	2
AGE-RAGE signaling pathway in diabetic complications	2
EGFR tyrosine kinase inhibitor resistance	2
Fructose and mannose metabolism	1
Glyoxylate and dicarboxylate metabolism	1
Propanoate metabolism	1
Inositol phosphate metabolism	1
Nitrogen metabolism	1
Fatty acid biosynthesis	1
Fatty acid elongation	1
Steroid biosynthesis	1
Steroid hormone biosynthesis	1
Ether lipid metabolism	1
Arachidonic acid metabolism	1
Linoleic acid metabolism	1
alpha-Linolenic acid metabolism	1
Pyrimidine metabolism	1
Alanine, aspartate and glutamate metabolism	1
Glycine, serine and threonine metabolism	1
Arginine and proline metabolism	1
Selenocompound metabolism	1
Cyanoamino acid metabolism	1
N-Glycan biosynthesis	1
Various types of N-glycan biosynthesis	1
Mucin type O-glycan biosynthesis	1
Glycosphingolipid biosynthesis - globo and isoglobo series	1
Thiamine metabolism	1
Folate biosynthesis	1
Retinol metabolism	1
Flavone and flavonol biosynthesis	1
Aflatoxin biosynthesis	1
Styrene degradation	1
Metabolism of xenobiotics by cytochrome P450	1

Drug metabolism - cytochrome P450	1
Basal transcription factors	1
Ribosome	1
Aminoacyl-tRNA biosynthesis	1
Nucleocytoplasmic transport	1
mRNA surveillance pathway	1
Nucleotide excision repair	1
Two-component system	1
MAPK signaling pathway - plant	1
ErbB signaling pathway	1
Hedgehog signaling pathway	1
VEGF signaling pathway	1
JAK-STAT signaling pathway	1
TNF signaling pathway	1
Calcium signaling pathway	1
Phospholipase D signaling pathway	1
cGMP-PKG signaling pathway	1
Cell adhesion molecules	1
Peroxisome	1
Apoptosis - fly	1
Ferroptosis	1
Tight junction	1
Hematopoietic cell lineage	1
Neutrophil extracellular trap formation	1
Toll-like receptor signaling pathway	1
NOD-like receptor signaling pathway	1
Cytosolic DNA-sensing pathway	1
Natural killer cell mediated cytotoxicity	1
B cell receptor signaling pathway	1
Fc epsilon RI signaling pathway	1
Leukocyte transendothelial migration	1
Chemokine signaling pathway	1
Regulation of lipolysis in adipocytes	1
Adipocytokine signaling pathway	1
GnRH secretion	1
Estrogen signaling pathway	1
Prolactin signaling pathway	1
Oxytocin signaling pathway	1
Relaxin signaling pathway	1
Aldosterone synthesis and secretion	1
Cortisol synthesis and secretion	1
Adrenergic signaling in cardiomyocytes	1
Vascular smooth muscle contraction	1
Carbohydrate digestion and absorption	1
Mineral absorption	1
Aldosterone-regulated sodium reabsorption	1
Cholinergic synapse	1
Long-term depression	1
Retrograde endocannabinoid signaling	1
Neurotrophin signaling pathway	1
Phototransduction	1
Osteoclast differentiation	1
Longevity regulating pathway - worm	1
Plant-pathogen interaction	1
Chemical carcinogenesis - DNA adducts	1
Central carbon metabolism in cancer	1
PD-L1 expression and PD-1 checkpoint pathway in cancer	1
Glioma	1

Melanoma	1
Endometrial cancer	1
Human T-cell leukemia virus 1 infection	1
Herpes simplex virus 1 infection	1
Vibrio cholerae infection	1
Epithelial cell signaling in Helicobacter pylori infection	1
Pathogenic Escherichia coli infection	1
Yersinia infection	1
Tuberculosis	1
Bacterial invasion of epithelial cells	1
Malaria	1
Chagas disease	1
Rheumatoid arthritis	1
Spinocerebellar ataxia	1
Morphine addiction	1
Dilated cardiomyopathy	1
Type I diabetes mellitus	1
Maturity onset diabetes of the young	1
Platinum drug resistance	1

Table S4. Statistics of transcriptome data at different developmental stages.

Sample name	Raw reads	Raw bases(G)	Q20(%)	Q30 (%)	GC content (%)	Clean reads	Unique and multiple genome mapping rate
Second_stage_larvae1	48,135,584	7.22	97.7	93.3	38.3	46,917,958	98.40%
Second_stage_larvae2	39,351,906	5.9	97.6	93.1	38.8	38,394,672	98.81%
Second_stage_larvae3	44,881,742	6.73	97.8	93.5	38.4	43,855,344	98.75%
Second_stage_larvae4	53,550,168	8.03	97.6	93.2	38.2	52,138,400	98.97%
Third_stage_larvae1	46,296,974	6.94	97.6	93	37.3	45,378,286	96.68%
Third_stage_larvae2	41,360,512	6.2	97.1	92	38.9	40,484,508	95.91%
Third_stage_larvae3	49,172,660	7.38	98.1	94.1	39.4	48,261,134	96.83%
Third_stage_larvae4	43,515,862	6.53	97.9	93.7	38.1	42,798,582	97.95%
Pupae1	44,369,424	6.66	97.6	93	37.4	43,269,180	97.72%
Pupae2	45,405,746	6.81	97.8	93.4	34.8	44,266,870	98.08%
Pupae3	45,369,684	6.81	97.5	92.8	37.1	44,230,024	96.21%
Pupae4	45,207,468	6.78	97.7	93.2	37.2	44,240,666	96.46%
Adult1	44,370,474	6.66	97.7	93.4	37.2	43,325,756	97.74%
Adult2	46,274,526	6.94	97.6	93	36	45,156,092	97.59%
Adult3	44,920,656	6.74	97.6	92.9	35.3	43,736,892	97.25%
Adult4	46,819,830	7.02	97.3	92.5	36.4	45,586,092	97.61%

6.3 Article 3

Jia, Z., Hasi, S., Zhan, D., Hou, B., Vogl, C., & Burger, P. A. (2023). Genome and transcriptome analyses facilitate genetic control of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly. *Insects*, 14(7), 620. https://doi.org/10.3390/insects14070620

Impact factor: 3.0

The data created within the article has been deposited at the following locations:
All transcriptome data were deposited at the National Center for Biotechnology Information's Sequence Read Archive (NCBI's SRA) database (https://www.ncbi.nlm.nih.gov/sra) under the BioProject PRJNA941182. The accession number for the Iso-Seq data of the mixed different developmental stages and sexes is SRR23730896. The accession numbers for RNA-seq data of adult females and adult males range from SRR23731231 to SRR23731236. Supplementary materials, including Figure S1, File S1 to File S3, Table S1, Table S4, and Table S5, can be found immediately following the main text of this article in this thesis. Supplementary materials Table S2 and Table S3 were deposited at Phaidra Vetmeduni with the identifier https://phaidra.vetmeduni.ac.at/o:2417.

insects

MDPI

Article

Genome and Transcriptome Analyses Facilitate Genetic Control of *Wohlfahrtia magnifica*, a Myiasis-Causing Flesh Fly

Zhipeng Jia [1], Surong Hasi [2], Deng Zhan [2], Bin Hou [2], Claus Vogl [3] and Pamela A. Burger [1],*

[1] Research Institute of Wildlife Ecology, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Savoyenstrasse 1, 1160 Vienna, Austria
[2] Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture and Rural Affairs, Inner Mongolia Agricultural University, Hohhot 010018, China
[3] Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Veterinaerplatz 1, 1210 Vienna, Austria
* Correspondence: pamela.burger@vetmeduni.ac.at

Simple Summary: *Wohlfahrtia magnifica*, a flesh fly, parasitizes several warm-blooded vertebrates and causes severe traumatic myiasis, detrimental to animal welfare and the livestock industry across Eastern and Southern Europe, Northern Africa, and Western and Northeast Asia. Genetic control has emerged as an effective and promising alternative to insecticides for controlling insect pests. In this study, we isolated and characterized two sex-determination genes, *W. magnifica transformer (Wmtra)* and *W. magnifica transformer2* (*Wmtra2*). These investigations may contribute to the establishment of genetically modified strains in *W. magnifica*. For example, the regulated first intron of *Wmtra*, a key component in the conditional female lethal transgenic systems, can be used to control the sex-specific expression of a pro-apoptotic gene, as developed for myiasis-causing blow flies, *Lucilia cuprina* and *Cochliomyia hominivorax*. Additionally, we performed a differential expression gene analysis between adult males and adult females and identified five candidate genes (*vasa* (*vas*), *nanos* (*nanos*), *bicoid* (*bcd*), *Bicaudal C* (*BicC*), and *innexin5* (*inx5*)) from the female-biased gene set that could upregulate Cas9 expression in the germline in Cas9-based homing gene drive systems, as established in mosquitoes. In summary, the isolation and characterization of these genes provide a solid foundation for the development of genetic control programs against *W. magnifica*.

Abstract: Myiasis caused by *Wohlfahrtia magnifica* is a widespread parasitic infestation in mammals. The infested host suffers from damage as the developing larvae feed on its tissues. For the control of myiasis infestation, genetic methods have been shown to be effective and promising as an alternative to insecticides. Combining genome, isoform sequencing (Iso-Seq), and RNA sequencing (RNA-seq) data, we isolated and characterized two sex-determination genes, *W. magnifica transformer (Wmtra)* and *W. magnifica transformer2* (*Wmtra2*), whose orthologs in a number of insect pests have been utilized to develop genetic control approaches. *Wmtra* transcripts are sex-specifically spliced; only the female transcript encodes a full-length functional protein, while the male transcript encodes a truncated and non-functional polypeptide due to the presence of the male-specific exon containing multiple in-frame stop codons. The existence of five predicted TRA/TRA2 binding sites in the male-specific exon and the surrounding intron of *Wmtra*, as well as the presence of an RNA-recognition motif in WmTRA2 may suggest the auto-regulation of *Wmtra* by its own protein interacting with WmTRA2. This results in the skipping of the male-specific exon and translation of the full-length functional protein only in females. Our comparative study in dipteran species showed that both the WmTRA and WmTRA2 proteins exhibit a high degree of similarity to their orthologs in the myiasis-causing blow flies. Additionally, transcriptome profiling performed between adult females and adult males reported 657 upregulated and 365 downregulated genes. Functional analysis showed that among upregulated genes those related to meiosis and mitosis Gene Ontology (GO) terms were enriched, while, among downregulated genes, those related to muscle cell development and aerobic metabolic processes were enriched. Among the female-biased gene set, we detected five candidate genes, *vasa* (*vas*), *nanos* (*nanos*), *bicoid* (*bcd*), *Bicaudal C* (*BicC*), and *innexin5* (*inx5*). The promoters of these genes

Citation: Jia, Z.; Hasi, S.; Zhan, D.; Hou, B.; Vogl, C.; Burger, P.A. Genome and Transcriptome Analyses Facilitate Genetic Control of *Wohlfahrtia magnifica*, a Myiasis-Causing Flesh Fly. *Insects* **2023**, *14*, 620. https://doi.org/10.3390/insects14070620

Academic Editor: Steve Whyard

Received: 31 May 2023
Revised: 29 June 2023
Accepted: 4 July 2023
Published: 10 July 2023

may be able to upregulate Cas9 expression in the germline in Cas9-based homing gene drive systems as established in some flies and mosquitoes. The isolation and characterization of these genes is an important step toward the development of genetic control programs against *W. magnifica* infestation.

Keywords: *Wohlfahrtia magnifica*; myiasis; *tra* gene; *tra2* gene; Iso-Seq; RNA-seq; genetic control

1. Introduction

Wohlfahrtia magnifica (Schiner, 1862; Diptera, Sarcophagidae) is an obligate parasitic species belonging to the group of flesh flies that cause severe myiasis in livestock, such as horses [1,2], sheep [3,4], camels [5,6], and even in humans [7,8]. Similar to other myiasis-causing flies, female adults of *W. magnifica* are attracted by wounds or natural body orifices of the host, such as the genitalia, and deposit the first instar larvae there. For subsequent development, the first- to third-stage larvae feed on the tissues, leading to serious health consequences for the host.

In regions where *W. magnifica* is distributed, from Eastern and Southern Europe and Northern Africa to Western and Northeast Asia [6,9–17], *W. magnifica*-related myiasis has led to important animal welfare and health problems, as well as huge economic losses due to reproduction problems, lameness, blindness, and even death if the infestation stays untreated [9,18,19]. As of now, a large number of cases of myiasis resulting from *W. magnifica* have been documented. For example, in Spain, Remesar et al. investigated a total of 73,683 sheep from 122 flocks in Albacete Province, and the results indicated the overall flock prevalence of traumatic myiasis was 95.9%, with an individual prevalence of 7.1% [20]; in China, Liu et al. surveyed 2038 female camels in selected sites from May to October 2021 in Inner Mongolia, and the results showed that the overall prevalence rate was 26.6% [21]. Killing the larvae with insecticides is the method most employed to fight myiasis-causing flies infestation. But frequent use of insecticides can result in resistance, necessitating an increase in the insecticide dosage until it eventually loses its efficacy. Furthermore, it is toxic to beneficial insects and non-target species in the local environment. Long-term prevention of *W. magnifica* and other myiasis-causing fly infestation is not reliably achieved using insecticides. For example, *Lucilia cuprina*, a myiasis-causing blow fly in Australia and New Zealand, has developed resistance to a wide range of insecticides by metabolic and target site insensitivity-resistance mechanisms [22].

Genetic control holds significant potential in effectively and promisingly managing insect pests. This approach aims to suppress the population size of target pest species to a non-critical level through targeting their reproductive capacity. Developed in the 1950s by Raymond Bushland, Edward Knipling, and colleagues [23,24], the sterile insect technique (SIT) is the best-known, as well as very successful, genetic control strategy. For example, *Cochliomyia hominivorax*, a blow fly that is an obligatory myiasis agent, has been successfully eradicated in North and Central America using SIT [25,26]. According to its guiding principles, SIT entails mass-rearing insects in special facilities, subjecting them to a high dosage of ionizing radiation, and dispersing them widely in predetermined regions. As a result, sterile male flies can mate with females from a wild population, resulting in no offspring being produced and further declines in insect populations over several generations. Traditionally, both sexes are released. However, the co-released females can compete with the wild females for mating with the released sterile males, which can increase the number of insects required for population suppression. In field tests with sterilized insect pests, such as the Mediterranean fruit fly in Guatemala, releasing male-only SIT may be three to five times more effective than the bisexual release in reducing the targeted populations [27].

Conditional female lethal transgenic strains for the myiasis-causing flies *C. hominivorax* [28] and *L. cuprina* [29] were produced considering the advantages of male-only releases and were found to be highly effective. The system consists of a driver construct expressing the *tetracycline transactivator* (*tTA*) gene under the control of a promoter, and an effector construct composed of a tTA-regulated pro-apoptotic gene, such as the *head involution defective* (*hid)* gene. When adding the antidote tetracycline to the diet, tetracycline can bind to tTA and thereby prevent the expression of the pro-apoptotic gene. In contrast, when tetracycline is absent, tTA can bind to a tetracycline operator (tetO), promoting the expression of the pro-apoptotic gene. To make the system sex-specific, the sex-specifically spliced first intron of *tra* is introduced within the pro-apoptotic gene. As a result, only females die when insects are reared with a diet lacking in tetracycline, while both females and males can survive with a diet containing tetracycline.

As another potential approach for insect pest control, Cas9-based homing gene drives have been established in a variety of pest species, in particular the mosquitoes *Anopheles gambiae*, *Anopheles stephensi*, and *Aedes aegypti* [30–34]. In its simplest form, the "homing construct" system consisting of a Cas9 nuclease and a guide RNA (gRNA) is designed to insert precisely into the genome. The Cas9 nuclease is guided by a gRNA to cleave a target site on the wild-type chromosome and form the double-strand break. Subsequently, taking the locus incorporating the homing construct as a template, the double-strand break can undergo homology-directed repair (HDR), a naturally occurring nucleic acid repair process. By copying similar sequences, this repair mechanism can result in the perfect copying of the drive allele containing the homing construct into the wild-type chromosome and effectively converting a heterozygote into a homozygote. Referred to as "super-Mendelian" inheritance, the frequency of transmitting the drive allele to the next generation is greater than expected by random segregation of heterozygous alleles, potentially enabling it to suppress pest populations. In a Cas9-based homing gene drive system, the selection of gene drive targets is essential. As the reproductive capacity of female flies determines the growth of insect populations, female development or reproduction genes could be outstanding candidates. For example, Carrami et al. generated a Cas9-based homing gene drive strain targeting the *tra* gene and showed its high efficiency for sex conversion from females to males in *D. melanogaster* [35].

In the sex determination pathway of some dipteran species, the TRA–TRA2 complex autoregulated the female-specific splicing of *tra* pre-mRNA and directed the splicing of the pre-mRNA of the transcription factor *doublesex* (*dsx*), whose protein, DSX, promotes sexual development by regulating the transcription of sex-specific differentiation genes. In addition, *tra* in the myiasis-causing blow flies, *L. cuprina* and *C. hominivorax*, has been used to create a conditional female lethal transgenic strain as it is responsible for turning sex-specific expression of a pro-apoptotic gene on or off. *Tra2* is often used as a target gene for genetic control strategies of insect pests. In this study, based on genome, isoform sequencing (Iso-Seq), and RNA sequencing (RNA-seq) data, we isolated and characterized two sex-determination genes, *W. magnifica transformer (Wmtra)* and *W. magnifica transformer2* (*Wmtra2*), with the aim of laying the foundation for the development of a conditional female lethal transgenic strain. From the female-biased gene set, we identified five candidate genes, *vasa* (*vas*), *nanos* (*nanos*), *bicoid* (*bcd*), *Bicaudal C* (*BicC*), and *innexin5* (*inx5*), whose promoters can drive Cas9 expression in the germline in Cas9-based homing gene drive systems, as established in some flies and mosquitoes. The isolation of these genes is an important step toward the development of genetic control programs for *W. magnifica* infestation.

2. Materials and Methods

2.1. Genome Resources of W. magnifica

In a previous publication [36], we reported sequencing, assembling, and annotating the genome of *W. magnifica*. The genome was deposited in GeneBank with accession number JAKWBJ000000000 under BioProject PRJNA778059. In addition, the annotation file and the putative transcripts and proteins of the genome of *W. magnifica* are available on Dryad (https://doi.org/10.5061/dryad.qfttdz0j8, accessed on 5 May 2022).

2.2. W. magnifica Sample Collection

In the study, the research species, *W. magnifica*, is an invertebrate agricultural insect pest, which is not an endangered or protected species. Second-stage and third-stage larvae samples of *W. magnifica* were collected non-invasively from domestic Bactrian camels in the field in Siziwang Banner, Ulanqab City, Inner Mongolia, China, therefore no animal experimental or ethical permits were necessary. The experimental protocols for the flies followed the procedures of Inner Mongolia Agricultural University. In short, the third-stage larvae were divided into two parts, one of which was placed in a foam box containing local soil; the rest along with the second-stage larvae were dropped directly into liquid nitrogen and then stored in a refrigerator at −80 °C. Subsequently, a portion of the three-day-old pupae was picked out of the soil in the foam box, and stored in a refrigerator at −80 °C. After 14 days, when the remaining pupae emerged into adult flies, the sex was distinguished, and adult females and adult males were placed into the refrigerator at −80 °C.

2.3. RNA Isolation and Assessment

The total RNA of each sample was extracted with the RNA Easy Fast Tissue/Cell kit (Tiangen Biotech, Beijing, China) following the manufacturer's instructions. The concentration, purity, and integrity of the extracted RNA were measured using NanoDrop (Thermo Fisher Scientific, Wilmington, DE, USA), Agilent 5400 (Agilent Technologies, Palo Alto, CA, USA), and 1% agarose gels. Qualified RNA samples were used for PacBio and Illumina library construction.

2.4. Illumina RNA-Seq Library Construction, Sequencing and Data Filtering

High-quality total RNA extracted from six samples, including three females and three males (each sample with one individual), was used for RNA-seq library preparation using the NEBNext® Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's instructions. In brief, the polyA fraction (mRNA) was purified from total RNA using oligonucleotides (dT) magnetic beads. The purified mRNA was fragmented and cDNA synthesized followed by end repair, A-tailing, adapter ligation, and PCR amplification steps. The prepared library was evaluated using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and qualified libraries were sequenced on an Illumina NovaSeq platform. Clean reads were generated by removing the adaptor sequences, low-quality reads, contamination from Bactrian camel and rRNA and by keeping reads with a minimum length of 75 base pairs (bp) using BBduk in the BBTools toolset [37].

2.5. PacBio Iso-Seq Library Construction, Sequencing, and Data Processing

For Iso-Seq, the total RNA of different developmental stages and sexes was pooled in equal amounts. Subsequently, mRNA was isolated and reverse-transcribed into full-length cDNA using the SMARTer PCR cDNA Synthesis Kit (Clontech, Palo Alto, CA, USA). Two SMRTbell libraries were constructed using the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, CA, USA). The prepared libraries were sequenced on the PacBio Sequel II platform.

Iso-Seq raw data stored in the BAM files were processed using the CCS program v6.4.0 (https://github.com/PacificBiosciences/ccs, accessed on 27 October 2022) with default parameters, and circular consensus sequences (CCS) were called. CCS reads containing the 5′ primer, the 3′ primer and, the polyA tail were processed by primer removal for generating the full-length (FL) reads using the lima program v2.6.0 (https://github.com/pacificbiosciences/barcoding/, accessed on 5 November 2022) with the parameters: --isoseq --dump-clips --peek-guess. Next, the refine module of the IsoSeq3 program v3.8.1 (https://github.com/PacificBiosciences/IsoSeq, accessed on 5 November 2022) was employed to identify and remove polyA tails and concatemers to generate full-length non-concatemer (FLNC) reads. FLNC reads were clustered to generate transcripts using the cluster module of the IsoSeq3 program v3.8.1. As a result, high-quality and low-quality isoforms were obtained.

The pbmm2 program v1.9.0 (https://github.com/PacificBiosciences/pbmm2, accessed on 5 November 2022), a minimap2 SMRT wrapper for PacBio Iso-Seq data, was applied to map high-quality isoforms onto the reference genome of *W. magnifica* [36]. With the mapping results, the redundant isoforms were collapsed using the collapse module of the IsoSeq3 program v3.8.1.

2.6. Isolation of the Wmtra and Wmtra2 Genes

We isolated the *Wmtra* and *Wmtra2* genes from the collapsed Iso-Seq transcript dataset. However, we did not obtain a full-length male transcript of the *Wmtra* gene, probably because the male-specific transcript is lower-expressed. Therefore, the male-specific transcript was reconstructed by aligning the three male RNA-seq data to the *W. magnifica* genome [36] using the HISAT2 program v2.2.1 [38] and feeding the output to the StringTie program v2.2.1 [39] for a genome-based assembly. The obtained *Wmtra* and *Wmtra2* transcripts were aligned to the *W. magnifica* genome [36] using the Minimap2 program v.2.24 [40] for genomic organization analysis.

2.7. Reverse Transcription Polymerase Chain Reaction (RT-PCR) Validation

The same batch of total RNA of both adult females and adult males with RNA-seq was used for RT-PCR. Based on the sequence obtained by Iso-Seq, we designed primers in 5′ untranslated region and the third common exon to amplify the sex-specific region using Primer3Plus (https://www.primer3plus.com/, accessed on 4 May 2023) and oligos were listed as follows:

Wmtra-F: 5′-CGGGAAGGTTAGGCTGTAGC-3′;

Wmtra-R: 5′-CGCAGATGAGGGTGGAGAAG-3′.

RT-PCR analysis for *Wmtra* was performed using the PrimeScript™ One Step RT-PCR Kit Ver.2 (Takara, Dalian, China), in which RNA→cDNA→PCR reactions were amplified in a single reaction system. Following the protocol's instruction, PrimeScript 1 Step Enzyme Mix, 2X 1 Step Buffer, WmtraF, WmtraR, total RNA, and RNase Free dH2O were added to a 50 µL reaction system to amplify the sex-specific region under the condition of 1 cycle of 50 °C for 30 min and 94 °C for 2 min; 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min. RT-PCR products were visualized on gel electrophoresis and then sent for Sanger sequencing.

Since we detected two sequences of *Wmtra2* in the collapsed Iso-Seq dataset, we designed three pairs of primers to investigate whether both sequences are verifiably transcribed in *W. magnifica* and whether *Wmtra2* is sex-specific. The primer pair 1 include the start and the stop codons or regions in their close proximity; the primer pair 2 was designed by moving outwards. As both sequences were identical, except that it was 129 bp longer at the 3′ terminal ends, we designed the reverse primer Wmtra2-R3 of the primer pair 3 within this fragment to verify whether the longer sequence was present in *W. magnifica*. The primers for the *Wmtra2* amplification were as follows:

Wmtra2-F1: 5′-ATGAGTCCTCGTTCACGCAG-3′;
Wmtra2-R1: 5′-ACTGACACACTTCAAGGGGC-3′;
Wmtra2-F2: 5′-ACGGCTTTGCTTTTGTACAGT-3′;
Wmtra2-R2: 5′-ATGCATATGGTTCGATGGAATAAAT-3′;
Wmtra2-F3: 5′-TGGCGAAATTGAACATTTACGGA-3′;
Wmtra2-R3: 5′-AATTTCTTTCAAGTCTTTATTTTGCCT-3′.

Since we did not obtain the expected *wmtra2* product using the one-step RT-PCR method, we amplified *wmtra2* using a two-step approach, where reverse transcription and PCR are reacted in separate tubes. PCR reaction conditions were set to 1 cycle of 98 °C for 2 min; 35 cycles of 98 °C for 20 s, 55 °C for 20 s and, 72 °C for 30 s; 1 cycle of 72 °C for 5 min; and 1 cycle of 16 °C for 2 min.

2.8. Sequence Analysis

A multiple alignment of protein sequences was performed using Clustal Omega [41]; the analysis of the alignment results was performed with Jalview v.2.11.2.6 [42]. Phylogenetic analysis was carried out using the neighbor-joining method in the MEGA program v.11.0.13 [43] with 1000 bootstrap replicates. Accession numbers for TRA sequence analysis used in this study are *Lucilia sericata* (AGE31795.1), *L. cuprina* (ACS34687), *C. hominivorax* (AGE31793.1), *Cochliomyia macellaria* (AGE31794.1), *Bactrocera oleae* (CAG29241.1), *Ceratitis capitata* (XP_004526947.1), *Musca domestica* (ACY40709.1), *Drosophila melanogaster* (AAF49441.1), and *Drosophila virilis* (EDW68645.2). Accession numbers for TRA2 sequence analysis used in this study include *L. cuprina* (ACS34688.1), *C. hominivorax* [44], *L. sericata* (XP_037815979.1), *C. capitata* (ACC68674.1), *M. domestica* (AAW34233.1), *B. oleae* (CAD67988.1), *D. melanogaster* (AAA28953.1), *Drosophila suzu*kii (ATI14861.1), *D. virilis* (EDW60892.2), *Stomoxys calcitrans* (NP_001298164.1), *Bactrocera correcta* (AJE26246.1), *Anastrepha bistrigata* (CBJ17289.1), and *Anastrepha obliqua* (CBJ17280.1).

2.9. Identification of Differentially Expressed Genes (DEGs)

We used three adult female and three adult male samples to investigate DEGs. The clean reads of each sample were mapped to the genome of *W. magnifica* [36] using the HISAT2 program v. 2.2.1 [38]. With the aligned bam files as input, raw counts of each sample were generated with the featureCounts program v2.0.3 [45]. In addition, raw count values were normalized the transcript per million (TPM). Prior to differential gene expression analysis, we also conducted a principal component analysis (PCA) after regularized log transformation (rlog) of TPM by the rlog function of the DESeq2 R package [46]. Subsequently, a differential expression analysis of genes was performed with the DESeq2 R package [46] using a q-value of < 0.05 and fold change ≥ 2 as a cutoff for the assignment of DEGs. GO enrichment analysis of DEGs was conducted with a cut-off criterion of q-value < 0.05.

2.10. Promoter Analysis

We extracted the upstream sequences of the start codon of *Wmnanos* by 2000 bases as a regulatory region harboring the promoter. The transcription start site and the putative TATA box were identified with BDGP (https://www.fruitfly.org/seq_tools/promoter.html, accessed on 19 May 2023). We used AliBaba2.1 (http://gene-regulation.com/pub/programs/alibaba2/, accessed on 19 May 2023) to predict transcription factor binding sites. AliBaba2.1 was set to the default settings except for Pairsim and Matrix conservation which were set to 64 and 80%, respectively.

2.11. Identification of Target Genes against W. magnifica Infestation

We followed the approach of Anstead et al. [47], who exploited functional genomic data of the extensively studied fruitfly *D. melanogaster* as a resource and inferred the functions of 988 genes of *L. cuprina*, whose orthologs in *D. melanogaster* were single-copy and associated with (semi-)lethality. In our study, we used the same 988 protein sequences in *D. melanogaster* as a query to search against the protein set of *W. magnifica* with the BLASTP program v2.7.1 (E-value $\leq 1 \times 10^{-20}$). If an ortholog of these proteins was detected in *W. magnifica*, we considered it as a potential target for the development of vaccines, drugs, or genetic control measures.

3. Results

3.1. Isolation and Characterization of the Wmtra Gene

Based on the Iso-Seq data, we successfully identified a female full-length transcript of the *Wmtra* gene of 1748 bp (Figure 1A and Supplementary Materials File S1). It consists of an open reading frame encoding 410 amino acids (Supplementary Materials File S1), as well as a 222 bp long 5′ untranslated region and a 293 bp long 3′ untranslated region. We also reconstructed (see Section 2.6) a male transcript of 2026 bp from the genome-based assembly of the male RNA-seq data (Figure 1A and Supplementary Materials File S1). The male transcript encodes a short protein of 63 amino acids (Supplementary Materials File S1), which is truncated and non-functional, because of the absence of the serine-arginine dipeptide-rich region (RS domain) involved in protein–protein interactions. Transcript differences of *Wmtra* between males and females result from a similar sex-specific splicing pattern (Figure 1A) as in the blow flies *C. hominivorax* [48] and *L. cuprina* [49].

The PCR verification results showed a 551 bp RT-PCR product in female flies, while in males we detected an 829 bp product, the extra 278 bp being the male-specific exon, which is consistent with the sequencing results (Figure 1D).

The *Wmtra* gene includes five exons and three introns (Figure 1A). The exons 1–4 are common in the transcripts of females and males, while the exon M1 is male-specific, containing multiple in-frame translation stop codons (Figure 1A). Except for different splice donor sites in the first intron, the splicing pattern between the male and female transcript is identical (Figure 1A,C). The exon M1 is located between the common exons 1 and 2 and is contiguous with the common exon 1 (Figure 1A).

Within the *Wmtra* sequence, five TRA/TRA2 binding sites were identified (Figure 1A,B). Among them, four clustered sites are located in the first intron and one in the exon M1.

The multiple alignments of protein sequences indicate that the first, second, and third introns occur at identical positions in *Wmtra*, *Lctra, Chtra*, and *Lstra* (Figure 2A). In addition, we found up to 50.26%, 50.95%, and 52.76% identities between the WmTRA and LcTRA, ChTRA, and LsTRA proteins, respectively. WmTRA contains a characteristic serine-arginine dipeptide-rich region (RS domain) and a proline-rich region at the C-terminal end (proline-rich domain) (Figure 2A). In addition, a TRACAM (C, *Ceratitis*; A, *Apis*; M, *Musca*) domain and a conserved DIP (DIPTERA) domain in dipteran species were identified (Figure 2A). The phylogenetic analysis shows that the WmTRA protein and Calliphoridae TRA proteins form a cluster and are more closely related to each other than to Muscidae, Tephritidae, or Drosophilidae species (Figure 2B).

Insects **2023**, *14*, 620 8 of 22

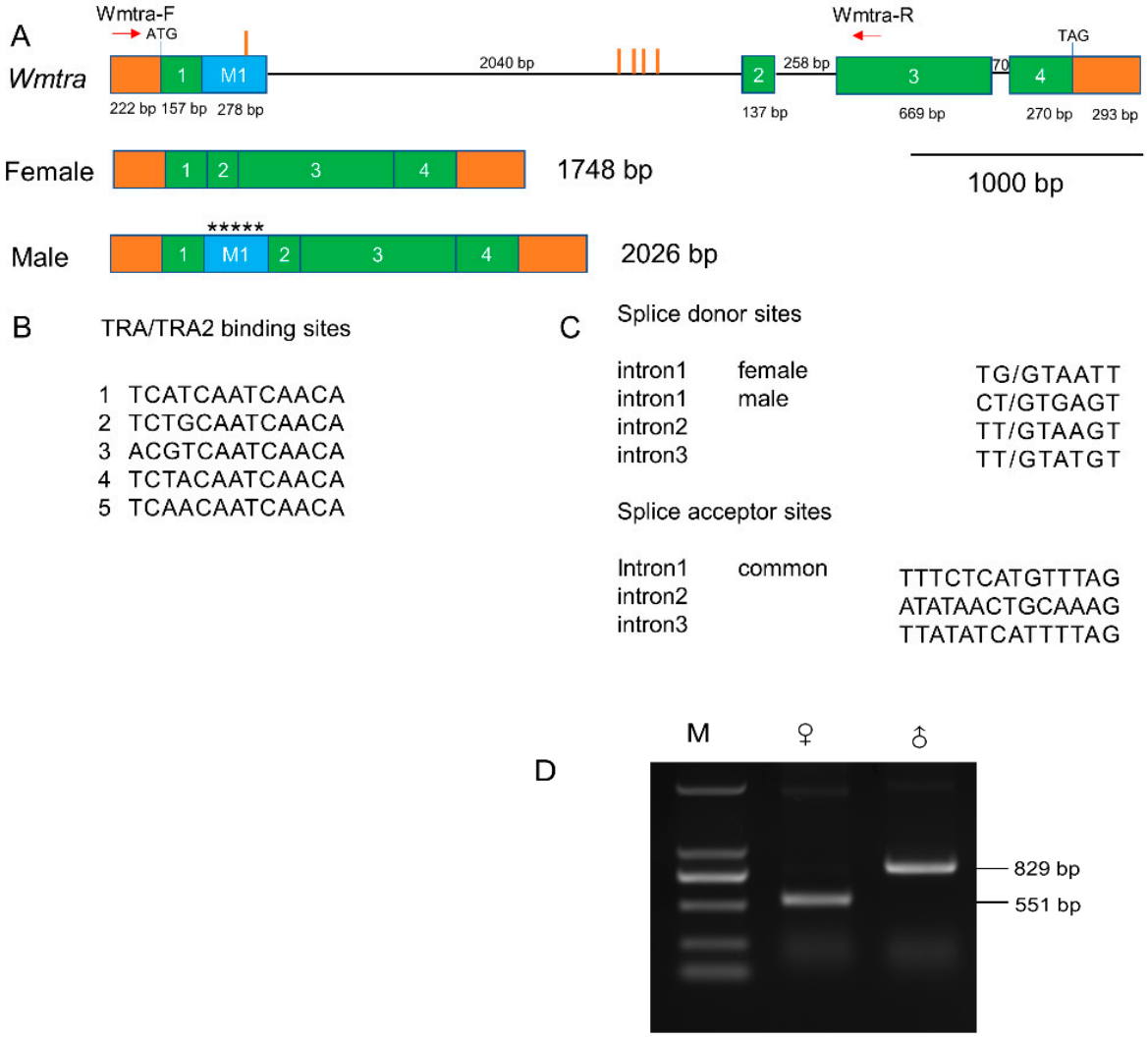

Figure 1. Genomic organization and sex-specific transcripts of *Wmtra*. (**A**) The *Wmtra* gene on the top diagram consists of four common exons (boxes 1, 2, 3, and 4 in both female and male transcripts (green boxes), and a male-specific exon M1 (blue box). Introns are represented by black horizontal lines. The 5′- and 3′-untranslated regions are shown in orange boxes. The translational start and stop locations are marked, and exon and intron lengths are shown in bp. The red vertical lines indicate the locations of the putative TRA/TRA2 binding sites in and around the exon M1. The red arrows indicate the primers. Transcripts for males and females are shown below the gene. Five asterisks in the exon M1 represent multiple in-frame stop codons (TAA, TGA, TAG, TAA, and TAG). (**B**) Sequences of the five TRA/TRA2 binding sites identified in the genomic DNA of the *Wmtra* gene. (**C**) Splice donor and acceptor sites of all *Wmtra* introns. The intron 1 "female" donor site is used to produce the female transcript and the intron 1 "male" donor site is used for males. The intron 1 "common" acceptor site is used to produce both female and male transcripts. (**D**) The detection of sex-specific transcripts of the *Wmtra* gene by RT-PCR analysis. M, ♀ and ♂ indicate the marker, adult females and adult males, respectively.

The multiple alignments of protein sequences indicate that the first, second, and third introns occur at identical positions in *Wmtra*, *Lctra*, *Chtra*, and *Lstra* (Figure 2A). In addition, we found up to 50.26%, 50.95%, and 52.76% identities between the WmTRA and LcTRA, ChTRA, and LsTRA proteins, respectively. WmTRA contains a characteristic serine-arginine dipeptide-rich region (RS domain) and a proline-rich region at the C-terminal end (proline-rich domain) (Figure 2A). In addition, a TRACAM (C, *Ceratitis*; A, *Apis*; M, *Musca*) domain and a conserved DIP (DIPTERA) domain in dipteran species were identi-

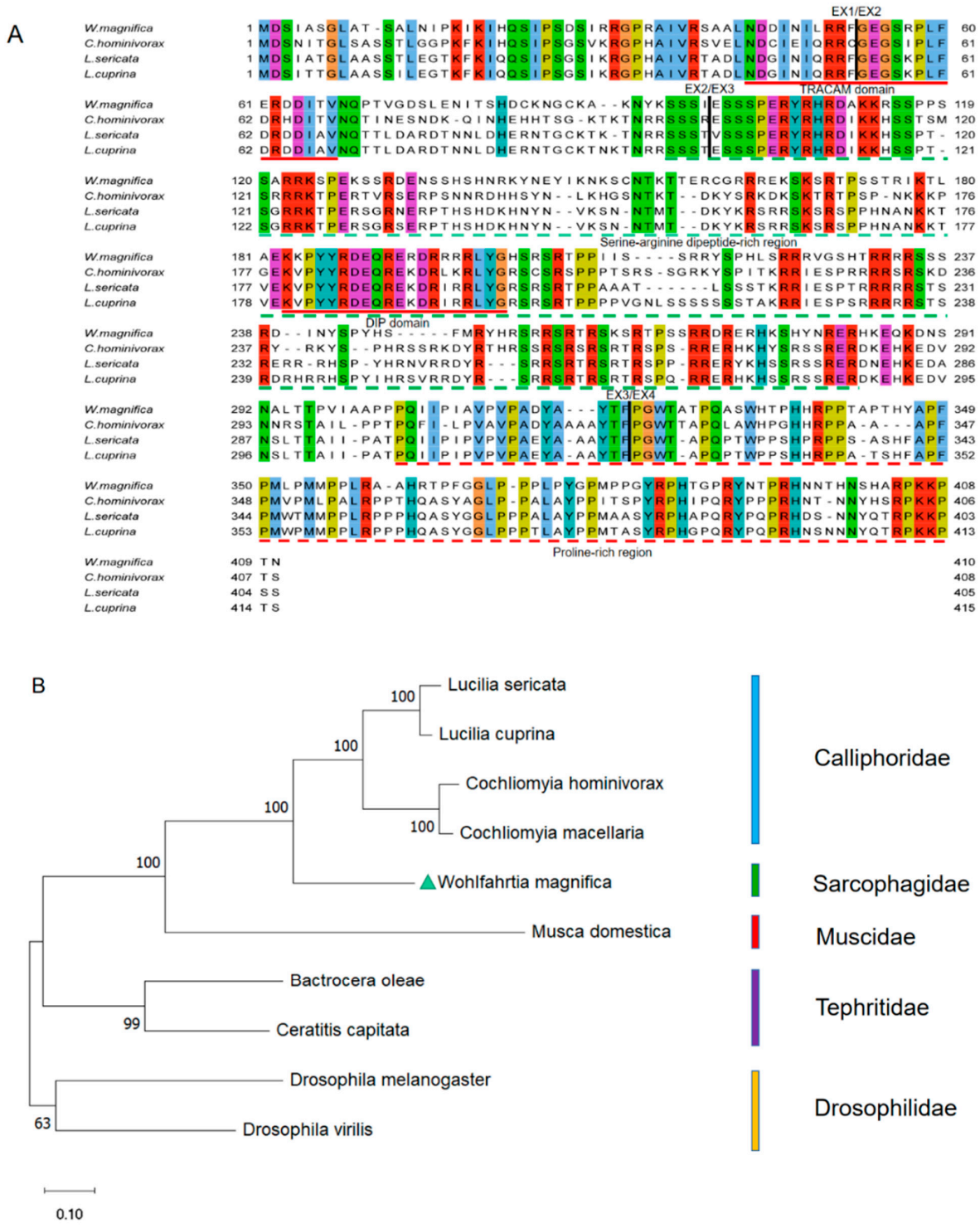

Figure 2. Multiple sequence alignments and phylogenetic analysis of TRA proteins among *W. magnifica* and other dipteran species. (**A**) Multiple sequence alignment of TRA proteins from *W. magnifica*, *C. hominivorax*, *L. sericata* and *L. cuprina*. Identical amino acids are shaded in the same color. The corresponding positions of the exon/intron boundaries are indicated in the TRA proteins by black vertical lines. The red horizontal lines represent the TRACAM domain and the DIP domain. The green and red horizontal dotted lines represent the serine-arginine dipeptide-rich region and the proline-rich region, respectively. (**B**) The neighbor-joining tree of selected TRA proteins from dipteran species. The numbers represent bootstrap support values from 1000 replicates. The green triangle highlights *W. magnifica*.

start codon is located at the last three bases of the first exon and the stop codon is in the 15th–17th bases of the eighth exon. Two sequences of *Wmtra2* with 1285 bp and 1414 bp in length (Supplementary Materials File S2) were found from the collapsed Iso-Seq dataset. Except for the difference in length of the 3′ terminal end, the other regions of the two sequences are identical, encoding a putative protein of 258 amino acids (Supplementary Materials File S2).

The PCR validation results showed that the amplification products using primer pair 1 and primer pair 2 were 798 and 953 bp in length, respectively, and the product lengths are consistent in male flies and female flies (Figure 3B). In contrast, using the primer pair 3, there was no PCR product (Figure 3B), suggesting that *Wmtra2* does not transcribe the 1414 bp long sequences, which may be a redundancy in the Iso-Seq dataset. This is consistent with *L. cuprina*, which transcribes a single non-sex-specific transcript.

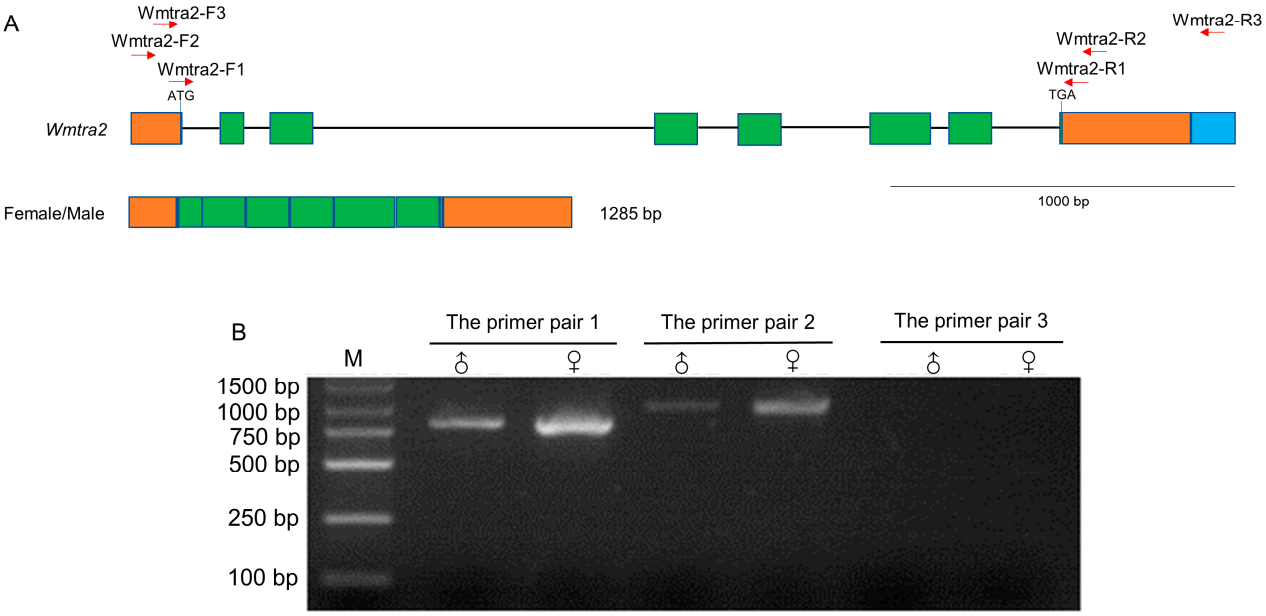

Figure 3. Genomic organization and the transcript of *Wmtra2*. (**A**) The *Wmtra2* gene on the top diagram consists of 8 exons. The green and orange regions represent coding regions and untranslated regions, respectively. Introns are represented by black horizontal lines. The blue region represents the difference between two sequences of *Wmtra2* from the Iso-Seq dataset. The translational start and stop locations are marked. The red arrows indicate the primers. The transcript for females/males is shown below the gene. (**B**) The detection of the transcript of the *Wmtra2* gene by RT-PCR analysis. M, P, and ♂ and ♀ indicate the marker, adult female and adult male, respectively.

The multiple alignment of protein sequences between the WmTRA2 protein and the TRA2 proteins from other myiasis-causing flies shows that the WmTRA2 protein contains an RNA-recognition motif (RRM) with two ribonucleoprotein regions (RNP1 and RNP2) immediately followed by the linker region and flanked by a serine-arginine dipeptide-rich N-terminal region (RS1 domain) and a serine-arginine dipeptide-rich C-terminal region (RS2 domain), which mediate protein–protein interactions (Figure 4A). The RS1 domain is mainly encoded by exons 2, 3, and 4, the RRM domain by exons 5 and 6, and the RS2 domain by exons 7 and 8. Phylogenetic analysis between the WmTRA2 protein and the TRA2 proteins from other dipteran species shows that, similarly to WmTRA, the WmTRA2 protein clusters with TRA2 proteins in Calliphoridae as these species belong to Oestroidea (Figure 4B).

domain by exons 7 and 8. Phylogenetic analysis between the WmTRA2 protein and the TRA2 proteins from other dipteran species shows that, similarly to WmTRA, the WmTRA2 protein clusters with TRA2 proteins in Calliphoridae as these species belong to Oestroidea (Figure 4B).

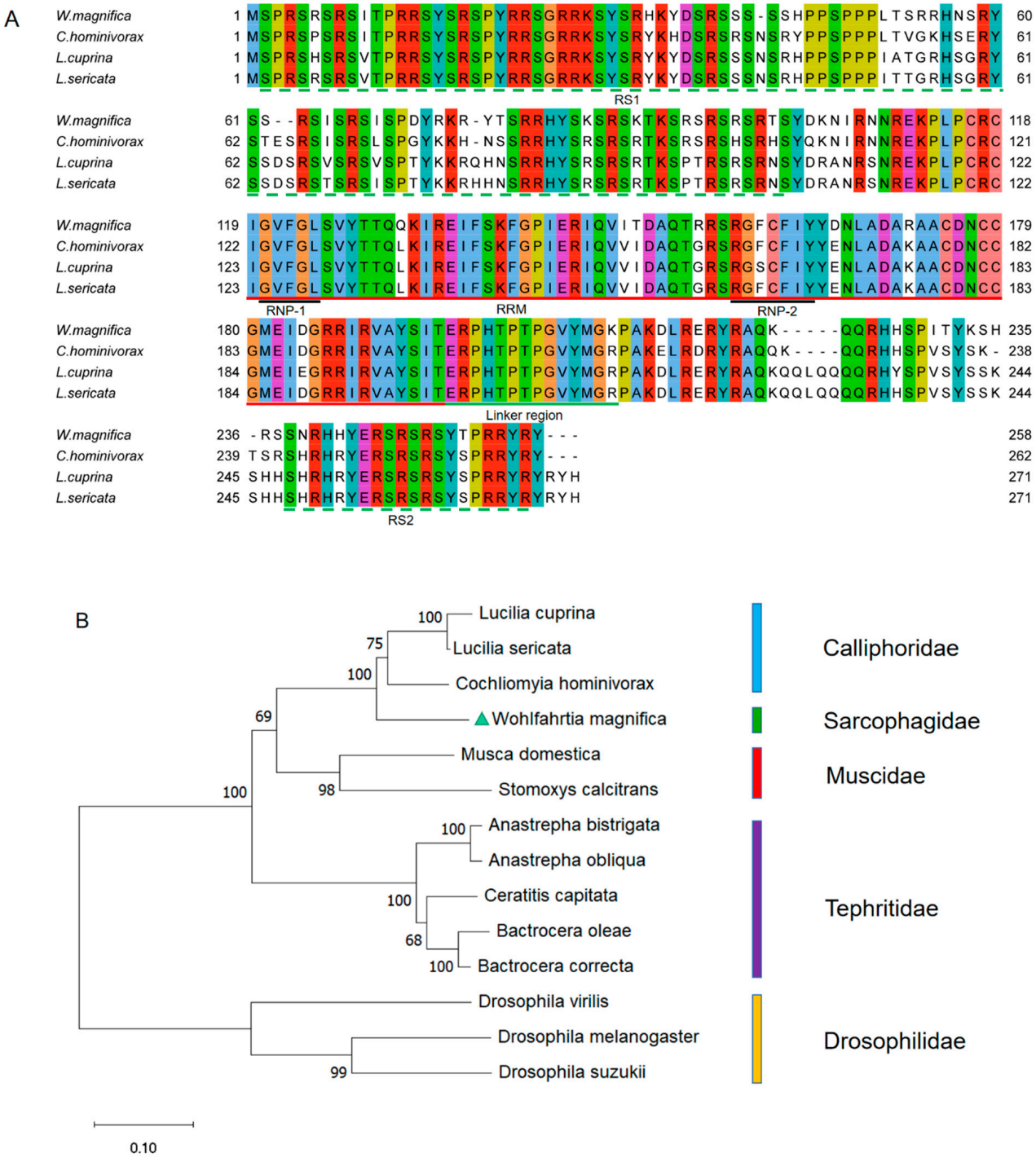

Figure 4. Multiple sequence alignments and phylogenetic analysis of TRA2 proteins among *W. magnifica* and other dipteran species. (**A**) Multiple sequence alignments of TRA2 proteins from *W. magnifica*, *C. hominivorax*, *L. cuprina*, and *L. sericata*. Identical amino acids are shaded in the same color. The red, black, and green horizontal lines represent RRM, RNP, and the linker region, respectively. The green horizontal dotted lines represent the RS1 domain and the RS2 domain. (**B**) The neighbor-joining tree of selected TRA2 proteins from dipteran species. The numbers represent bootstrap support values from 1000 replicates. The green triangle highlights *W. magnifica*.

Insects **2023**, *14*, 620 12 of 22

3.3. Gene Expression Analysis between Adult Females and Adult Males

After removing low-quality reads, adaptors, rRNA, and contaminants, a total of 38.66 Gb clean paired-end reads, including 18.44 Gb from females and 20.22 Gb from males, were used for downstream analyses (Supplementary Materials Table S1). The number of clean reads per sample varies between 39,186,716 and 47,109,424 among the six sequenced samples (Supplementary Materials Table S1). The PCA results show that the biological replicates of male and female samples are distributed in two separate groups (Supplementary Materials Figure S1). In addition, more than 92% of clean reads from each sample can be mapped to the *W. magnifica* genome (Supplementary Materials Table S1).

Between adult males and adult females, 1022 genes were found to be differentially expressed, of which 365 were downregulated and 657 upregulated in females (Figure 5 and Supplementary Materials Table S2). The top 10 terms of a subsequent GO enrichment analysis using DEGs are shown in Figure 6. We noted that genes upregulated in males are annotated for GO terms involved in muscle and mitochondrial structure (e.g., Cellular Component: "striated muscle thin filament" and "mitochondrial membrane") and in muscle cell development and aerobic metabolic processes (e.g., Biological Process: "striated muscle cell development", "muscle contraction", "striated muscle cell differentiation", and aerobic respiration) (Supplementary Materials Table S3). On the other hand, genes upregulated in females are annotated for GO terms involved in mitosis and meiosis (e.g., Cellular Component: "spindle midzone" and "condensed chromosome"; Biological Process: "chromosome segregation" and "nuclear division") (Supplementary Materials Table S3).

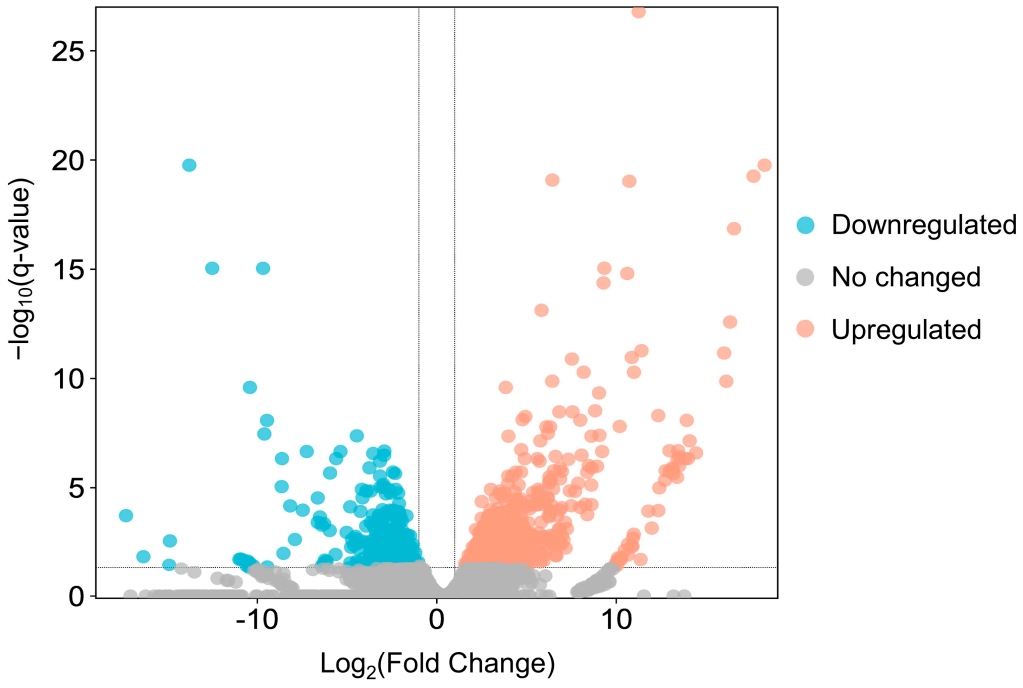

Figure 5. The volcano plot of DEGs in adult females versus adult males. Red dots and blue dots indicate upregulated genes and downregulated genes, respectively; grey dots are genes whose expression levels do not reach statistical significance.

Insects **2023**, *14*, 620 13 of 22

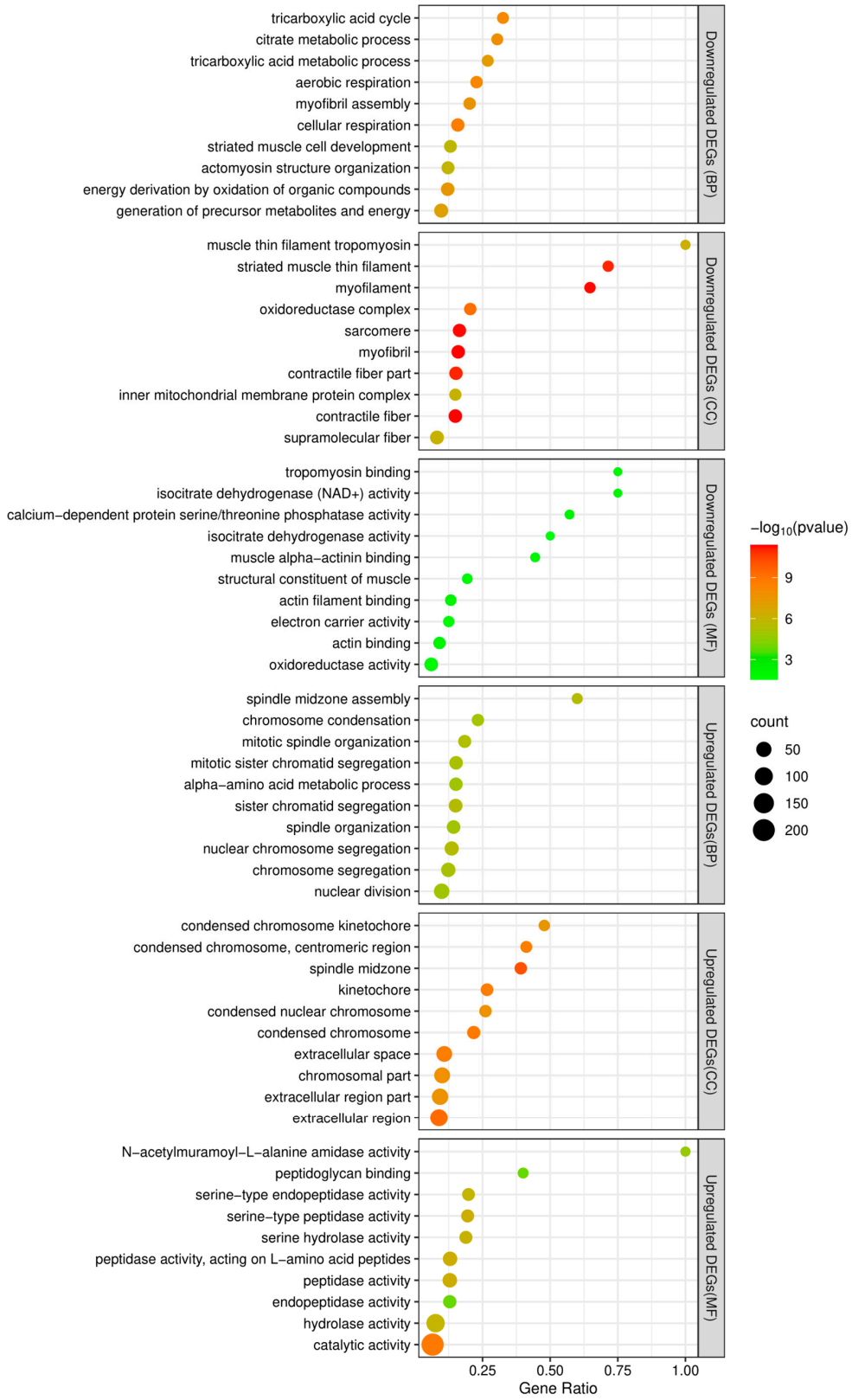

Figure 6. The top 10 enriched GO terms of functional enrichment analyses of downregulated and upregulated DEGs. The x-axis indicates the gene ratio and the y-axis represents the different GO terms. BP, CC, and MF represent Biological Process, Cellular Component, and Molecular Function groups of GO, respectively.

3.4. Candidate Genes for Cas9-Based Homing Gene Drive

From the female-biased gene set, several maternally expressed genes important for fertility (*inx5*) or embryonic development (*vas*, *nos*, *bcd*, and *BicC*) were identified, whose promoters could be used to drive Cas9 expression in the germline in Cas9-based homing gene drive systems.

vas (Woma_00004829) is localized at contig ctg.000023F. At approximately 6000 bp upstream of the start codon, we identified an ortholog of the *vasa intronic gene* (*vig*) of Drosophila, which was named SERPINE1 mRNA binding protein 1 (*Serbp1*, Woma_00004830) in *W. magnifica*. In *Drosophila*, the *vig* gene is located between the first non-coding exon and the downstream exon containing the start codon of the *vas* gene. Similarly, in *W. magnifica*, RNA-seq data supports that the *Wmvas* gene also initiates coding from the second exon, and the first non-coding exon does exist. This information is useful to accurately identify the sequence of the promoter. As the current annotation of the *W. magnifica* genome neglected the untranslated regions, the first non-coding exon of *Wmvas* is not annotated (Figure 7A). Similar to *C. hominivorax* and *L. sericata*, *Wmnanos* (Woma_00005371) contains four exons and the length of the coding region is 2622 bp. It is also linked to its upstream gene (Woma_00005370), the ortholog of CG11779 of *Drosophila* (Figure 7B). *bcd* (Woma_00011211) is present on contig ctg.000125F, and the organization is relatively simple with four exons (Figure 7C). *BicC* (Woma_00012414) is located in contig ctg.0000165F. *BicC* is relatively complicated with 12 exons (Figure 7D). Similar to other genes, *BicC* is only expressed in adult females, but *BicC* is not classified as a DEG. In *W. magnifica*, we identified an ortholog of zero population growth (*zpg*) (known as *inx4*), however, the expression levels of *Wmzpg* did not differ between adult males and adult females. In *C. hominivorax* [44] and *L. sericata* [50], no ortholog of *zpg* was found, and the *inx5* gene was considered to be an ortholog of *zpg*. Similarly, we also found an *inx5* gene (Woma_00010258) in contig ctg.000090F, which is very closely linked to *nudel* (Woma_00010257) and has only three exons (Figure 7E). The expression level analysis showed that all these genes, including *vas*, *nanos*, *bcd*, *BicC*, and *inx5*, were dominantly abundant in adult females (Figure 8 and Supplementary Materials Table S4).

We retrieved 2000 bp sequences upstream of the start codon of *Wmnanos* and performed a promoter analysis in silico. As a result, BDGP found a promoter sequence with a score of 1 containing the predicted transcription start site and TATA box (Figure 9A). Supported by AliBaba2.1, 49 transcription factor binding sites, such as for GATA binding protein (GATA-1), CCAAT/enhancer binding protein (C/EBP), activator protein-1 (AP-1), octamer-binding transcription factor-1 (Oct-1), Hunchback (Hb), TATA-binding protein (TBP), etc., were predicted (Figure 9B).

Furthermore, we identified six U6 RNA genes with 96 bp, 111 bp, 104 bp, 99 bp, 96 bp, and 106 bp in five contigs of the *W. magnifica* genome. Two U6 genes were present in contig ctg.000034F, while the other four U6 genes were distributed in contigs ctg.000025F, ctg.000033F, ctg.000037F, and ctg.000437F (Supplementary Materials File S3).

3.5. Potential Target Genes for Control Strategies against W. magnifica

In *W. magnifica*, we inferred 972 genes, whose *D. melanogaster* orthologs are single-copy and associated with lethality and semi-lethality upon disruption (Supplementary Materials Table S5). These genes can be used for screening potential candidate targets for the development of vaccines, insecticides, or genetic control measures.

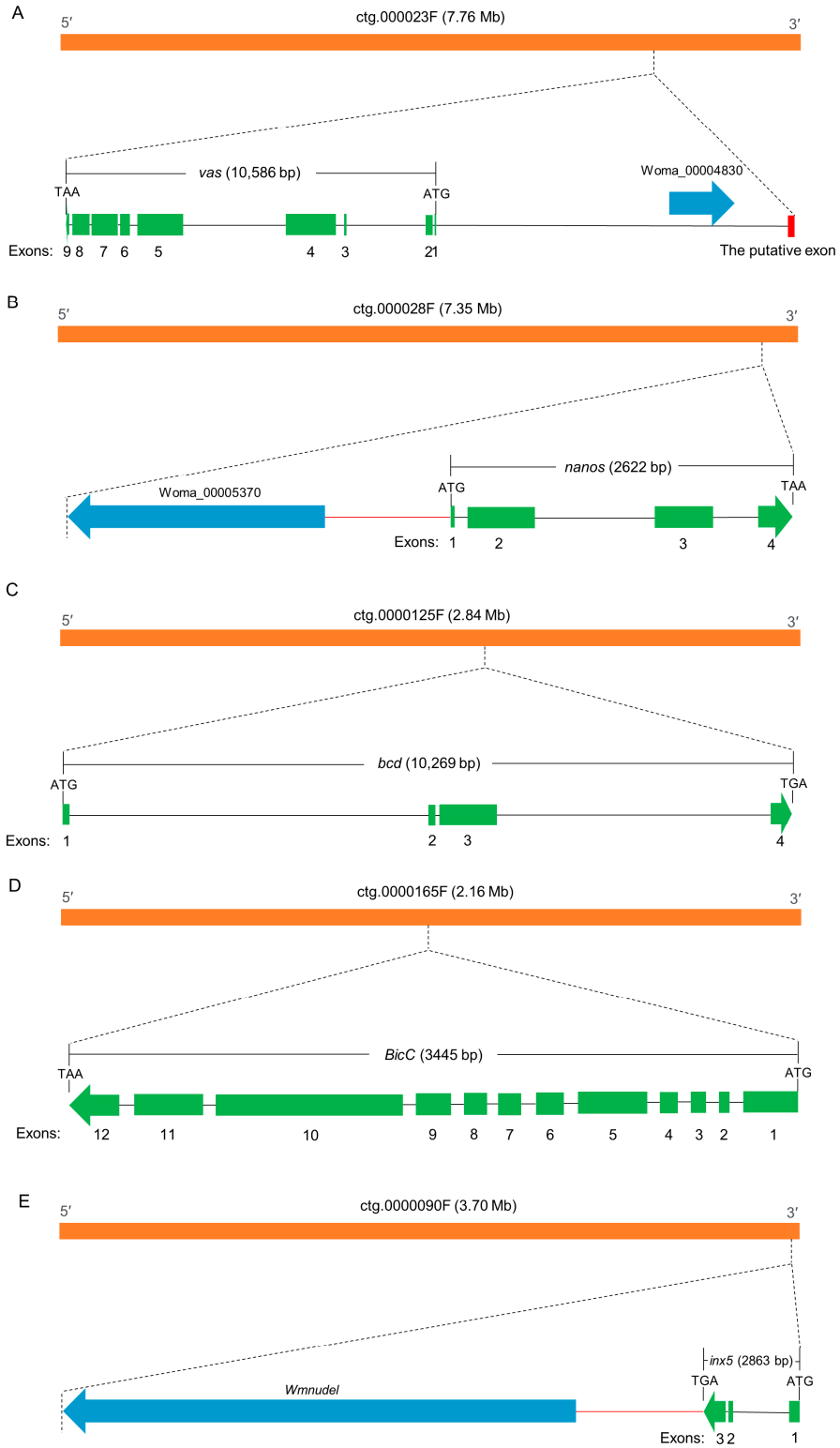

Figure 7. Diagrams of the genomic structures of *vas* (**A**), *nanos* (**B**), *bcd* (**C**), *BicC* (**D**), and *inx5* (**E**) in *W. magnifica*. The orange, green, blue, and red boxes represent contigs, exons, neighboring genes and the putative exon, respectively. The start and stop codons are marked. Arrows point in a 5′–3′ direction. The black horizontal line between the start codon and the stop codon indicates the introns. The red horizontal lines indicate intergenic regions, regulatory sequences, or/and untranslated regions. (Note: The *Wmvas* gene has two transcripts in the current genome annotation; in this figure, we only included the longest one).

W. magnifica. The orange, green, blue, and red boxes represent contigs, exons, neighboring genes and the putative exon, respectively. The start and stop codons are marked. Arrows point in a 5′–3′ direction. The black horizontal line between the start codon and the stop codon indicates the introns. The red horizontal lines indicate intergenic regions, regulatory sequences, or/and untranslated regions. (Note: The *Wmtra* gene has two transcripts in the current genome annotation; in this figure, we only included the longest one.)

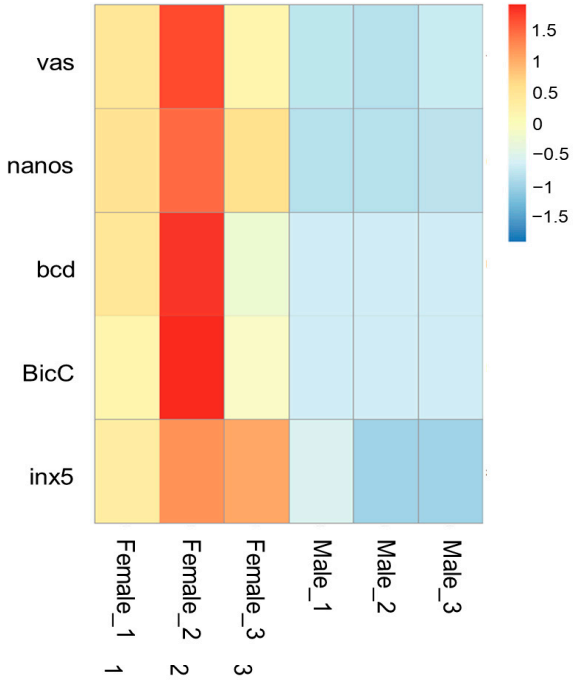

Figure 8. The heatmap of expression levels (RNA-seq analysis) of the *vas*, *nanos*, *bcd*, *BicC*, and *inx5* genes in females and males of *W. magnifica*. Rows represent genes and columns represent sample names. The red, yellow, and blue colors indicate high, medium, and low expressions, respectively.

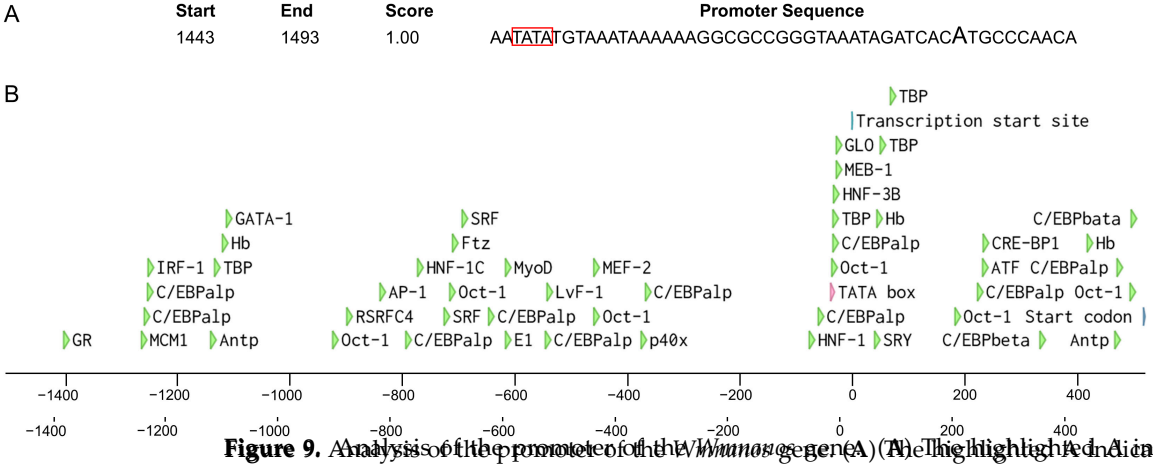

Figure 9. Analysis of the promoter of the *Wmmanos* gene. (**A**) The highlighted A indicates the predicted transcription start site. The TATA box is shown in red bold. Start and End represent the start and end of the sequence containing the transcription start site and the TATA box and the counting starts from the first position of the extracted 2000-bp sequence. Score represents the promoter score (between 0 and 1). (**B**) A schematic diagram of putative regulatory motifs in the 5′ flanking sequence. The colorful arrows represent the positions of transcription factor binding sites, the transcription start site or the start codon. The scale indicates the distance in nucleotides from the TSS at position 1.

4. Discussion

4.1. Acquisition of Full-Length Sequences of Transcripts

We obtained the female transcript of *Wmtra*, and the non-sex-specific transcript of *Wmtra2* using the Iso-Seq technique. Unexpectedly, we were unable to obtain the male transcript of *Wmtra* from the Iso-Seq dataset. Combining RNA-seq and Iso-Seq data, however, we found that most of the genes captured by Iso-Seq have relatively high expression, indicating that the obstacle to identifying the male-specific transcript of *Wmtra* may have been an insufficient Iso-Seq sequencing depth. Therefore, with the aim of obtaining more thorough transcripts, especially those with low relative expression in insects, it may be

necessary to conduct deeper sequencing. Usually, the full-length sequence of an RNA transcript can be obtained using a molecular biology approach known as rapid amplification of cDNA ends (RACE) if the sequence is only partially known. To date, many transcript sequences of *tra* or *tra2* in insects have been identified by the RACE method [48,49]. Iso-Seq represents an alternative method that enables the acquisition of full-length transcript sequences, including the entire coding sequence and untranslated regions. When combined with a technology that selects full-length capped and polyadenylated RNA molecules, the Iso-Seq method can maximize the repertoire of full-length transcripts for the objectives of the study. In comparison, the RACE approach only enables a limited number of target genes.

4.2. The Putative Sex Determination Mechanism of W. magnifica

The *tra and tra2* genes are pivotally important in insect sex determination. In the study, we isolated and characterized the two sex-determination genes in *W. magnifica*, *Wmtra* and *Wmtra2*, as well as their corresponding proteins, female WmTRA and WmTRA2. Similar to *L. cuprina* and *C. hominivorax*, *Wmtra* produces sex-specific transcripts and *Wmtra2* generates a single non-sex-specific transcript. We performed a multiple alignment of protein sequences between WmTRA/WmTRA2 and their orthologs in the myiasis-causing blow flies and the results indicate that TRA/TRA2 proteins are highly conserved among them. The phylogenetic analysis was in agreement with the taxonomic relationship, forming a cluster with proteins in blow flies.

In WmTRA, we also found four known TRA-specific domains [51–53] and two characteristic regions of the SR protein, including an RS domain and a proline-rich region at the C-terminal end [51,54]. The RS domain is found to mediate protein–protein interactions [54]. The second domain, TRACAM, was complete in the female WmTRA protein, but truncated in the male non-functional WmTRA protein due to the presence of the male-specific exon M1. In *M. domestica*, the molecular role of the TRACAM domain of MdTRA is in connection to the auto-regulatory function of *Mdtra* [55]. In *D. melanogaster*, the TRACAM domain is absent in non-auto-regulatory DmTRA. In contrast, a replaced *Sex-lethal* (*Sxl*) gene acts as an upstream regulator [52] instead of *tra*, suggesting the TRACAM domain may function in *Wmtra* auto-regulation. The other domains are found to be conserved in dipteran species, but the function remains unknown, such as the third domain (DIP domain) [51,55]. Furthermore, with a similar relative location to *L. cuprina* and *C. hominivorax*, we observed five TRA/TRA2 binding sites present in the male-specific exon and in the intron 1 of *Wmtra*, as well as two RS domains and an RNA-recognition motif in WmTRA2. These findings may indicate that WmTRA and WmTRA2 interact to form the TRA/TRA2 complex and bind to its own pre-mRNA, resulting in the auto-regulative splicing of *Wmtra* and the skipping of the male-specific exon. In non-Drosophilidae species, such as the medfly and the housefly, maternal deposition of *tra* mRNA in developing XX embryos translate into functional proteins and initiates the positive auto-regulatory loop of female-specific splicing, resulting in female differentiation [51,56]. In contrast, an M factor on the Y chromosome in XY embryos suppresses the *tra* function and, as a consequence, the initiation of the auto-regulatory loop is inhibited, promoting male development [56]. Future work investigating the expression pattern of *Wmtra* at different developmental stages, especially the early embryo, and the molecular function of the upstream Y-linked M factor will facilitate a better understanding of the sex determination mechanism in *W. magnifica*. Despite the successful identification of *tra* and *tra2* orthologs in *W. magnifica*, the investigation of gene function could not be pursued due to existing constraints on laboratory rearing of *W. magnifica*. Our next studies aim to employ advanced functional genomic tools, such as RNAi-mediated gene knockdown and CRISPR/Cas9-mediated gene knockout, for further investigation of their functions.

4.3. Genetic Controls against W. magnifica Infestation

Eradication of the New World screw-worm fly from the United States and later from Mexico and Central America through successive releases of radiation-sterilized flies pro-

duced at a mass-rearing facility demonstrates the effectiveness of SIT in insect pest control [25,26]. For SIT, however, females may consume half of the feed in a mass-rearing plant, although only males are effective in suppressing local populations. Eliminating females from the rearing process, therefore can result in significant savings in food costs. In a conditional female lethal transgenic strain developed for myiasis-causing blow flies by Concha et al. [28] and Yan et al. [29], only males survived in the absence of tetracycline. As an essential component to turn the sex-specific expression of a pro-apoptotic gene on or off in this system, the isolation of the *tra* gene plays an integral role in the development of the strain. In this study, we successfully identified and characterized *Wmtra*. Similar to other myiasis-causing flies, such as *L. cuprina*, *C. hominivorax*, *Wmtra* transcripts are spliced in a sex-specific manner, so that only the female transcript creates a full-length functional protein, whereas males encode, presumably, non-functional peptide. Therefore, by introducing the key first intron of the *Wmtra* gene to control the sex-specific expression of a pro-apoptotic gene in this system, we expect that the conditional female lethal transgenic system may work well in *W. magnifica*, as was the case for other myiasis-causing blow flies, *C. hominivorax* and *L. cuprina*, which has proven to be quite successful [28,29].

Gene drive systems promise to offer another powerful pest genetic control tool. Cas9-based homing gene drive systems have been developed in mosquitos [30–34]. For a successful gene drive system, the identification of precise sites in the target insect genome that are vital for female development, survival, or fecundity, is a key prerequisite. Furthermore, promoters from genes active in the germline to drive Cas9 expression, as well as promoters from U6 RNA for expression of gRNA are important. In the Medfly, *tra*-knockdown XY males develop normally, while XX individuals develop as fertile males [56]. Therefore, *Wmtra* can serve as such a target and female individuals are expected to be converted into males. Theoretically, this approach could result in an all-male population.

We analyzed DEGs between adult females and adult males and functional analysis showed that among upregulated genes those related to meiosis and mitosis GO terms were enriched, while among downregulated genes those related to muscle cell development and aerobic metabolic processes were enriched. These results likely correspond to the biology of the species, females invest in producing eggs, males in muscle for a mating flight. The female- and male-biased gene sets identified can also provide a useful resource for Cas9-based homing gene drive systems. For example, from the female-biased gene set, we identified several candidate genes, including *vas*, *nanos*, *bcd*, *BicC*, and *inx5*. Specifically, promoters from *nanos* and *vas* have been used in gene drive strains of mosquitoes to direct expression of the Cas9 nuclease [32]. In *D. suzukii*, promoters from either early germ cells, e.g., *vas* or *nanos*, or from late germ cells, e.g., *BicC*, have successfully driven the expression of the Cas9 nuclease [57]. Moreover, in the synthetic Medea toxin-antidote gene drive system, the promoter from *bcd* was employed to express a maternal toxin [58]. Among these, the *nanos* gene with a relatively simple organization should facilitate the isolation of the promoter. Also, the promoter from *nanos* was successfully applied in *D. melanogaster* and *D. suzukii* [59,60]. Therefore, a promoter from the *nanos* is a suitable candidate for driving Cas9 expression for further development of Cas9-based homing gene drives of *W. magnifica*. We investigated the promoter region of *Wmnanos* and identified 49 transcription factor binding sites. In the position of the TATA box, TBP binding sites are present, which may indicate that TBP binds to the TBP motif, facilitates the assembly of the pre-initiation complex, and promotes the recruitment of other transcription factors and RNA polymerase, ultimately leading to the initiation of transcription.

In the present investigation, our research findings revealed the maternal expression patterns in these genes in accordance with previous observations in other dipteran species. Nevertheless, due to the oviparous nature of *W. magnifica* and the current challenges in establishing comprehensive laboratory-rearing protocols, acquiring early-stage embryos directly from the field remains exceptionally arduous. In prospective investigations, we will undertake comprehensive examinations and refinement of laboratory-rearing protocols for

W. magnifica, with the objective of detecting the gene expression patterns at early embryonic stages, facilitating the development of the genetically modified system in this species.

The promoter from the U6 RNA gene is ideal for driving gRNA transcription in Cas9-based homing gene drive systems. In different mosquito species, U6 regulatory sequences were employed to promote gRNA expression, with various degrees of activity [61]. In *W. magnifica*, the promoters from six U6 genes could be used to drive gRNA expression.

In summary, the identification of these genes can contribute to the genetic control of *W. magnifica*. In the conditional female lethal transgenic systems, the sex-specifically spliced *Wmtra* gene can be used to control the sex-specific expression of a pro-apoptotic gene. In the Cas9-based homing gene drive systems, *Wmtra* can serve as a target to convert females into males, and the promoters from *vas*, *nanos*, *bcd*, *BicC*, and *inx5*, as well as the promoters from U6 genes, can be applied to express Cas9 nuclease and gRNA in the germline, respectively.

5. Conclusions

We successfully isolated and characterized the sex-determining genes *Wmtra* and *Wmtra2* in *W. magnifica*. *Wmtra* transcripts are sex-specifically spliced so that only the female transcript encodes a full-length functional protein, while the male transcript encodes a truncated and non-functional polypeptide due to the presence of the male-specific exon M1 containing multiple in-frame stop codons. The existence of five putative TRA/TRA2 binding sites in and around the male-specific exon M1 of *Wmtra* and the presence of an RNA-recognition motif in WmTRA2 may suggest that WmTRA interacts with its own pre-mRNA through WmTRA-2, resulting in the skipping of the male-specific exon M1. The comparative study showed that both the WmTRA and WmTRA2 proteins exhibited a high degree of similarity to their orthologs in the myiasis-causing blow flies, *L. sericata*, *L. cuprina*, and *C. hominivorax*. The sex transcriptome analysis reported 657 upregulated and 365 downregulated genes. Functional analysis showed that upregulated genes related to meiosis and mitosis were enriched, while downregulated genes were enriched in muscle cell development and aerobic metabolic processes. From the female-specific gene set, we identified five candidate genes, *vas*, *nanos*, *bcd*, *BicC*, and *inx5*, whose promoters can drive Cas9 expression in the germline in Cas9-based homing gene drive systems, as established in some dipteran species. The identification and characterization of these genes represent an important step in the development of genetic control programs for *W. magnifica* infestation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/insects14070620/s1, Figure S1: The PCA results of RNA-seq of three adult females and three adult males. File S1: Gene, transcript, and protein sequences of *Wmtra*. File S2: Gene, transcript, and protein sequences of *Wmtra2*. File S3: U6 RNA sequences in *W. magnifica*. Table S1: Statistics of RNA-seq data of adult females and adult males. Table S2: DEGs between adult females and adult males. Table S3: GO enrichment analysis of DEGs. Table S4: Expression levels of five female-biased genes. Table S5: *W. magnifica* orthologs of essential genes in *D. melanogaster*.

Author Contributions: Conceptualization, Z.J. and P.A.B.; methodology, Z.J. and P.A.B.; validation, Z.J. and B.H.; data curation, Z.J.; resources, S.H. and D.Z.; writing—original draft preparation, Z.J. and P.A.B.; writing-revision and editing, all authors; supervision, P.A.B., C.V. and S.H.; funding acquisition, Z.J., P.A.B. and S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Chinese Scholarship Council (CSC, the number 201909150004), the National Natural Science Foundation of China (the number 32060815), and the Austrian Science Funds (FWF) project (the number P29623-B25).

Data Availability Statement: All transcriptome data were deposited at National Center for Biotechnology Information's Sequence Read Archive (NCBI's SRA) database (https://www.ncbi.nlm.nih.gov/sra, accessed on 8 March 2023) under the BioProject accession number PRJNA941182. The accession number for Iso-Seq data of the mixed different developmental stages and sexes is SRR23730896.

The accession numbers for RNA-seq data of adult females and adult males are from SRR23731231 to SRR23731236.

Acknowledgments: We gratefully acknowledge Professor Surong Hasi's laboratory members as well as the herdsmen of Siziwang Banner for their assistance in sample collection. We thank Barbara Wallner, Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna, for advice on primer design. Open Access Funding by the University of Veterinary Medicine Vienna.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farkas, R.; Képes, G.Y. Traumatic myiasis of horses caused by *Wohlfahrtia magnifica*. *Acta Vet. Hung.* **2001**, *49*, 311–318. [CrossRef] [PubMed]
2. Yan, L.; Zhang, M.; Tang, L.; Ente, M.; Ma, X.; Chu, H.; Li, K.; Hu, D.; Zhang, D. First reports of nasal and traumatic myiasis infection in endangered Przewalski's horses (*Equus ferus przewalskii*). *Int. J. Parasitol. Parasites Wildl.* **2019**, *9*, 21–24. [CrossRef] [PubMed]
3. Dehghani, R.; Zarghi, I.; Sayyedi, H.R. Genital myiasis of a sheep by *Wohlfahrtia magnifica*, in Ghamsar, Kashan, Iran. *Bangladesh J. Med. Sci.* **2014**, *13*, 332–335. [CrossRef]
4. Farkas, R.; Hall, M.J.R.; Daniel, M.; Börzsönyi, L. Efficacy of ivermectin and moxidectin injection against larvae of Wohlfahrtia magnifica (Diptera: Sarcophagidae) in sheep. *Parasitol. Res.* **1996**, *82*, 82–86. [CrossRef]
5. Moshaverinia, A.; Moghaddas, E.; Maleki, M.; Borji, H. Gingival myiasis of camel (Camelus dromedarius) caused by Wohlfahrtia magnifica. *Sci. Parasitol.* **2013**, *14*, 85–87.
6. Valentin, A.; Baumann, M.P.O.; Schein, E.; Bajanbileg, S. Genital myiasis (*Wohlfahrtiosis*) in camel herds of Mongolia. *Vet. Parasitol.* **1997**, *73*, 335–346. [CrossRef]
7. Çiftçiŏglu, N.; Altintaş, K.; Haberal, M. A case of human orotracheal myiasis caused by Wohlfahrtia magnifica. *Parasitol. Res.* **1996**, *83*, 34–36. [CrossRef]
8. Kokcam, I.; Saki, C.E. A case of cutaneous myiasis caused by *Wohlfahrtia magnifica*. *J. Dermatol.* **2005**, *32*, 459–463. [CrossRef]
9. Farkas, R.; Hall, M.J.R.; Kelemen, F. Wound myiasis of sheep in Hungary. *Vet. Parasitol.* **1997**, *69*, 133–144. [CrossRef]
10. Gaglio, G.; Brianti, E.; Abbene, S.; Giannetto, S. Genital myiasis by *Wohlfahrtia magnifica* (Diptera, Sarcophagidae) in Sicily (Italy). *Parasitol. Res.* **2011**, *109*, 1471–1474. [CrossRef]
11. Giangaspero, A.; Traversa, D.; Trentini, R.; Scala, A.; Otranto, D. Traumatic myiasis by *Wohlfahrtia magnifica* in Italy. *Vet. Parasitol.* **2011**, *175*, 109–112. [CrossRef] [PubMed]
12. Hall, M.J.R.; Testa, J.M.; Smith, L.; Adams, Z.J.O.; Khallaayoune, K.; Sotiraki, S.; Stefanakis, A.; Farkas, R.; Ready, P.D. Molecular genetic analysis of populations of Wohlfahrt's wound myiasis fly, Wohlfahrtia magnifica, in outbreak populations from Greece and Morocco. *Med. Vet. Entomol.* **2009**, *23*, 72–79. [CrossRef] [PubMed]
13. Ruiz Martínez, I.; Leclercq, M. Data on distribution of screwworm fly Wohlfahrtia magnifica (Schiner) in southwestern Europe (Diptera: Sarcophagidae). *Notes Fauniques Gembloux* **1994**, *28*, 53–60.
14. Sotiraki, S.; Farkas, R.; Hall, M.J.R. Fleshflies in the flesh: Epidemiology, population genetics and control of outbreaks of traumatic myiasis in the Mediterranean Basin. *Vet. Parasitol.* **2010**, *174*, 12–18. [CrossRef] [PubMed]
15. Yasuda, M. Morphology of the Larva of Wohlfahrtia magnifica Schin. Found in a Wound on a Camel in Inner Mongolia. *J. Chosen Nat. Hist. Soc.* **1940**, *7*, 27–36.
16. Farkas, R.; Hall, M.J.R.; Bouzagou, A.K.; Lhor, Y.; Khallaayoune, K. Traumatic myiasis in dogs caused by Wohlfahrtia magnifica and its importance in the epidemiology of wohlfahrtiosis of livestock. *Med. Vet. Entomol.* **2009**, *23*, 80–85. [CrossRef]
17. Diakakis, N.; Papadopoulos, E.; Hall, M.J.R.; Desiris, A. Post-traumatic complication due to Wohlfahrtia magnifica larvae on a horse. *Vet. Rec.* **2006**, *158*, 170–172. [CrossRef]
18. İpek, D.N.S.; Şaki, C.E.; Çay, M. The investigation of lipid peroxidation, anti-oxidant levels and some hematological parameters in sheep naturally infested with *Wohlfahrtia magnifica* larvae. *Vet. Parasitol.* **2012**, *187*, 112–118. [CrossRef]
19. Martinez, R.I.; Cruz, S.M.D.; Rodriguez, R.; Lopez, D.M.; Parra, M.S.; Navio, F.A. Myiasis caused by wohlfahrt/a magnjfica in southern spain. *Isr. J. Vet. Med.* **1987**, *43*, 34–41.
20. Remesar, S.; Otero, J.L.; Panadero, R.; Díez-Baños, P.; Díaz, P.; García-Díos, D.; Martínez-Calabuig, N.; Morrondo, M.P.; López, C. Traumatic myiasis by *Wohlfahrtia* magnifica in sheep flocks from southeastern Spain: Prevalence and risk factors. *Med. Vet. Entomol.* **2022**, *36*, 30–37. [CrossRef]
21. Liu, J.; Hou, B.; Wuen, J.; Jiang, N.; Gao, T.; Hasi, S. Epidemiological Investigation on Genital Myiasis of Bactrian Camels in Parts of Inner Mongolia, China. *J. Camel Pract. Res.* **2022**, *29*, 229–235. [CrossRef]
22. Sandeman, R.M.; Levot, G.W.; Heath, A.C.G.; James, P.J.; Greeff, J.C.; Scott, M.J.; Batterhamg, P.; Bowles, V.M. Control of the sheep blowfly in Australia and New Zealand–are we there yet? *Int. J. Parasitol.* **2014**, *44*, 879–891. [CrossRef] [PubMed]
23. Bushland, R.C.; Lindquist, A.W.; Knipling, E.F. Eradication of screw-worms through release of sterilized males. *Science* **1955**, *122*, 287–288. [CrossRef]
24. Knipling, E.F. Possibilities of insect control or eradication through the use of sexually sterile males. *J. Econ. Entomol.* **1955**, *48*, 459–462. [CrossRef]

25. Wyss, J.H. Screwworm eradication in the Americas. *Ann. N. Y. Acad. Sci.* **2006**, *916*, 186–193. [CrossRef]
26. Scott, M.J.; Concha, C.; Welch, J.B.; Phillips, P.L.; Skoda, S.R. Review of research advances in the screwworm eradication program over the past 25 years. *Entomol. Exp. Appl.* **2017**, *164*, 226–236. [CrossRef]
27. Rendón, P.; McInnis, D.; Lance, D.; Stewart, J. Medfly (Diptera: Tephritidae) genetic sexing: Large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *J. Econ. Entomol.* **2004**, *97*, 1547–1553. [CrossRef]
28. Concha, C.; Yan, Y.; Arp, A.; Quilarque, E.; Sagel, A.; de León, A.P.; McMillan, W.O.; Skoda, S.; Scott, M.J. An early female lethal system of the New World screwworm, Cochliomyia hominivorax, for biotechnology-enhanced SIT. *BMC Genet.* **2020**, *21*, 143. [CrossRef]
29. Yan, Y.; Scott, M.J. Building a transgenic sexing strain for genetic control of the Australian sheep blow fly Lucilia cuprina using two lethal effectors. *BMC Genet.* **2020**, *21*, 141. [CrossRef]
30. Kyrou, K.; Hammond, A.M.; Galizi, R.; Kranjc, N.; Burt, A.; Beaghton, A.K.; Nolan, T.; Crisanti, A. A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nat. Biotechnol.* **2018**, *36*, 1062–1066. [CrossRef]
31. Hammond, A.; Galizi, R.; Kyrou, K.; Simoni, A.; Siniscalchi, C.; Katsanos, D.; Gribble, M.; Baker, D.; Marois, E.; Russell, S.; et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nat. Biotechnol.* **2016**, *34*, 78–83. [CrossRef] [PubMed]
32. Carballar-Lejarazú, R.; Ogaugwu, C.; Tushar, T.; Kelsey, A.; Pham, T.B.; Murphy, J.; Schmidt, H.; Lee, Y.; Lanzaro, G.C.; James, A.A. Next-generation gene drive for population modification of the malaria vector mosquito, Anopheles gambiae. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22805–22814. [CrossRef] [PubMed]
33. Gantz, V.M.; Jasinskiene, N.; Tatarenkova, O.; Fazekas, A.; Macias, V.M.; Bier, E.; James, A.A. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E6736–E6743. [CrossRef] [PubMed]
34. Reid, W.; Williams, A.E.; Sanchez-Vargas, I.; Lin, J.; Juncu, R.; Olson, K.E.; Franz, A.W. Assessing single-locus CRISPR/Cas9-based gene drive variants in the mosquito Aedes aegypti via single-generation crosses and modeling. *G3-Genes Genom Genet.* **2022**, *12*, jkac280. [CrossRef] [PubMed]
35. Carrami, E.M.; Eckermann, K.N.; Ahmed, H.M.; Sánchez, C.H.M.; Dippel, S.; Marshall, J.M.; Wimmer, E.A. Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6189–6194. [CrossRef]
36. Jia, Z.; Hasi, S.; Vogl, C.; Burger, P.A. Genomic insights into evolution and control of Wohlfahrtia magnifica, a widely distributed myiasis-causing fly of warm-blooded vertebrates. *Mol. Ecol. Resour.* **2022**, *22*, 2744–2757. [CrossRef]
37. Bushnell, B. *BBTools: A Suite of Fast, Multithreaded Bioinformatics Tools Designed for Analysis of DNA and RNA Sequence Data*; Joint Genome Institute: Berkeley, CA, USA, 2018.
38. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [CrossRef]
39. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]
40. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [CrossRef]
41. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef]
42. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [CrossRef] [PubMed]
43. Kumar, S.; Tamura, K.; Nei, M. MEGA: Molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics* **1994**, *10*, 189–191. [CrossRef]
44. Scott, M.J.; Benoit, J.B.; Davis, R.J.; Bailey, S.T.; Varga, V.; Martinson, E.O.; Hickner, P.V.; Syed, Z.; Cardoso, G.A.; Torres, T.T.; et al. Genomic analyses of a livestock pest, the New World screwworm, find potential targets for genetic control programs. *Commun. Biol.* **2020**, *3*, 1–14. [CrossRef]
45. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [CrossRef] [PubMed]
46. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21. [CrossRef] [PubMed]
47. Anstead, C.A.; Korhonen, P.K.; Young, N.D.; Hall, R.S.; Jex, A.R.; Murali, S.C.; Hughes, D.S.T.; LEE, S.F.; Perry, T.; Stroehlein, A.J.; et al. Lucilia cuprina genome unlocks parasitic fly biology to underpin future interventions. *Nat. Commun.* **2015**, *6*, 1–11. [CrossRef]
48. Li, F.; Vensko, S.P.; Belikoff, E.J.; Scott, M.J. Conservation and sex-specific splicing of the transformer gene in the Calliphorids Cochliomyia hominivorax, Cochliomyia macellaria and Lucilia sericata. *PLoS ONE* **2013**, *8*, e56303. [CrossRef]
49. Concha, C.; Scott, M.J. Sexual development in Lucilia cuprina (Diptera, Calliphoridae) is controlled by the transformer gene. *Genetics* **2009**, *182*, 785–798. [CrossRef]

50. Davis, R.J.; Belikoff, E.J.; Dickey, A.N.; Scholl, E.H.; Benoit, J.B.; Scott, M.J. Genome and transcriptome sequencing of the green bottle fly, Lucilia sericata, reveals underlying factors of sheep flystrike and maggot debridement therapy. *Genomics* **2021**, *113*, 3978–3988. [CrossRef]
51. Hediger, M.; Henggeler, C.; Meier, N.; Perez, R.; Saccone, G.; Bopp, D. Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. *Genetics* **2010**, *184*, 155–170. [CrossRef]
52. Verhulst, E.C.; van de Zande, L.; Beukeboom, L.W. Insect sex determination: It all evolves around transformer. *Curr. Opin. Genet. Dev.* **2010**, *20*, 376–383. [CrossRef] [PubMed]
53. Geuverink, E.; Beukeboom, L.W. Phylogenetic distribution and evolutionary dynamics of the sex determination genes doublesex and transformer in insects. *Sex Dev.* **2014**, *8*, 38–49. [CrossRef] [PubMed]
54. Manley, J.L.; Tacke, R. SR proteins and splicing control. *Genes Dev.* **1996**, *10*, 1569–1579. [CrossRef]
55. Tanaka, A.; Aoki, F.; Suzuki, M.G. Conserved domains in the transformer protein act complementary to regulate sex-specific splicing of its own pre-mRNA. *Sex Dev.* **2018**, *12*, 180–190. [CrossRef]
56. Pane, A.; Salvemini, M.; Bovi, P.D.; Polito, C.; Saccone, G. The transformer gene in Ceratitis capitata provides a genetic basis for selecting and remembering the sexual fate. *Development* **2002**, *129*, 3715–3725. [CrossRef]
57. Kandul, N.P.; Belikoff, E.J.; Liu, J.; Buchman, A.; Li, F.; Yamamoto, A.; Yang, T.; Shriner, I.; Scott, M.J.; Akbari, O.S. Genetically Encoded CRISPR Components Yield Efficient Gene Editing in the Invasive Pest *Drosophila suzukii*. *CRISPR J.* **2021**, *4*, 739–751. [CrossRef] [PubMed]
58. Chen, C.H.; Huang, H.; Ward, C.M.; Su, J.T.; Schaeffer, L.V.; Guo, M.; Hay, B.A. A synthetic maternal-effect selfish genetic element drives population replacement in Drosophila. *Science* **2007**, *316*, 597–600. [CrossRef]
59. Champer, J.; Yang, E.; Lee, E.; Liu, J.; Clark, A.G.; Messer, P.W. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads through a cage population. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24377–24383. [CrossRef]
60. Yadav, A.K.; Butler, C.; Yamamoto, A.; Patil, A.A.; Lloyd, A.L.; Scott, M.J. CRISPR/Cas9-based split homing gene drive targeting doublesex for population suppression of the global fruit pest *Drosophila suzukii*. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2301525120. [CrossRef]
61. Bottino-Rojas, V.; James, A.A. Use of Insect Promoters in Genetic Engineering to Control Mosquito-Borne Diseases. *Biomolecules* **2023**, *13*, 16. [CrossRef]

Supplementary Materials for:

Genome and transcriptome analyses facilitate genetic control of *Wohlfahrtia magnifica*, a myiasis-causing flesh fly

Zhipeng Jia[1], Surong Hasi[2], Deng Zhan[2], Bin Hou[2], Claus Vogl[3], Pamela A. Burger[1*]

1. Research Institute of Wildlife Ecology, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Savoyenstrasse 1, 1160 Vienna, Austria
2. Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture and Rural Affairs, Inner Mongolia Agricultural University, Hohhot 010018, China
3. Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria
*Corresponding author: pamela.burger@vetmeduni.ac.at

The following contents are included:

Supplementary Figure S1
Supplementary Files S1 to S3
Supplementary Tables S1, S4, and S5

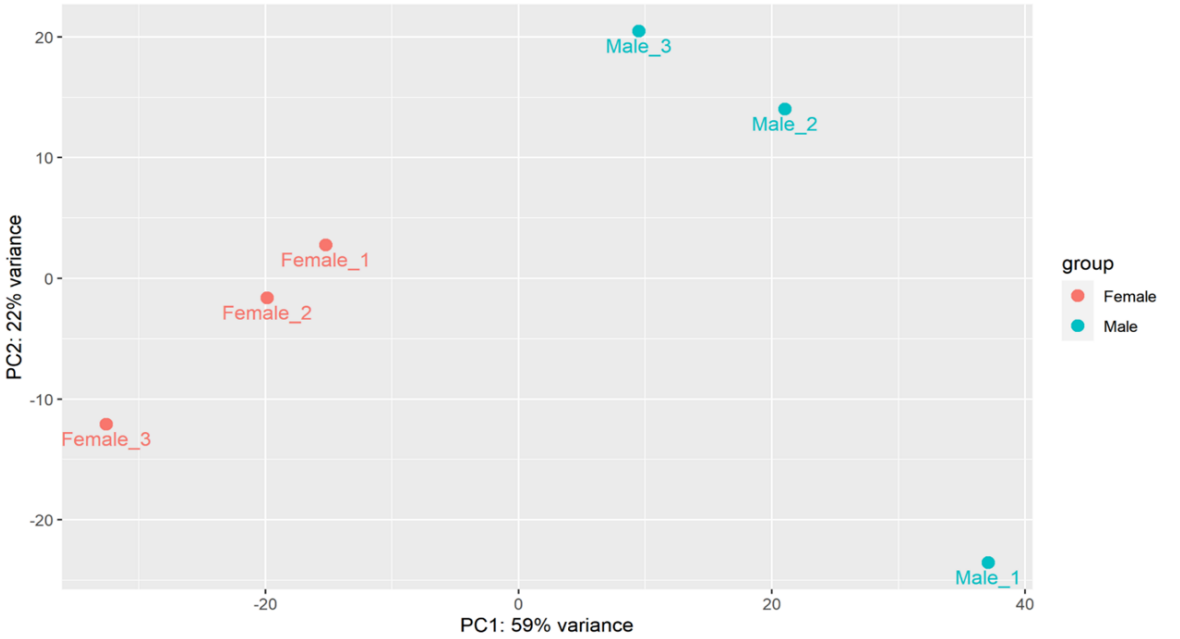

Figure S1. The PCA results of 3 adult females and 3 adult males. The red and blue dots represent adult females and adult males, respectively.

File S1. Gene, transcript, and protein sequences of *Wmtra*.

>ID=Woma_00009986; type=gene; loc=ctg.000086F:complement(833397..837790); name=*Wmtra*; length=4394

GGCGTCAAGTTGATTAAGAGTGTAGCAAAATTTTCGGGAAGGTTAGGCTGTAGCTAAAAGTACAAATTGGATAAAATTGGAGAC
CTTGGCGACATTGTTGTTTTTATTACAAAACAAATAAATATTTAGTTTTTTGGGCCAAAAGGAAAGTTTCGCCTGTTACCATCAAC
ACCATCATCTATCGCCTTCTCAACTTTATTACGACCACAGCAACAATTAGAAATGGATTCCATAGCATCTGGCTTAGCAACATCA
GCTCTAAATATACCAAAGATTAAAATTCATCAATCTATACCAAGTGATTCCATAAGAAGAGGTCCCCGTGCCATCGTACGCTCAG
CTGCTTTAAACGACGACATTAATATTTTACGCAGATTTGGTAATTTTCTTACAAAATTTCTGCATACCTTTTAAATTTTTTTGCAGT
ATTTAATACAATGAGAGTAGTTACAAAGAAATTAACAACATTATAAGGAACGTAAAAGAAATTAGATATTTTTATCAGTAATTAT
TGAATGGCACAATGAAAGAAAAAAGCTACTATTGAAGGCTGTGACGCGATATGCAGCGCTAGAAAAATTCATCAATCAACATTC
CATATCCATTTCCTTGAACAGTGAGGAGAATCTTCAAACAACTAAACCGCTGACTACAACTGTGAGTATTACAATTACATTTTTAT
TATATTAAAAAAAAAAAAACAATTTTCTAATGGTTCTAAGTGTTAACAATCGTTTTTCTTTTGTATATATTTCAAAAAATTCATTTTG
AAACAATCGCCTCAATGGTACACTCATACCTCCTTATACACCCTTTGTTTTTATGGCATTTCGCTATAACGGCTTGTCTAAATTGT
GGAACATCTTCGCTTTACGGCACAAAATACCCCATCATACGGTAAAAAAGACAGTATTTCATCGTTTTTAAGGAATAGTGTAAGA
CAAGCGAAATCAACGCGCTTTACGGCATTTTAATTTACGGCCTACATTTTCAGAGCCTAACTATGCCGCATAAAGAGGTGTGAGT
GCATTAAAAACAACCATTGTTGCTAGTGTCTATCGGGGTGGCATTTTTACTCTATACACCGATAAAGCTAATGTTATGGTATACAT
ATGTCCAGCTATACCCTAGTTTTTTTGTATTCACTACTTATTATCTTATTCTTAGTAATGCTGTGAAATACTTTTTAATTCTTTGTTC
GTCACTAAAGACAGTAATTACGTAGTAATTGATTTACTAAAATAATTGCTCTATCATTTGGTAACTAGCACTCCTACTGCAATTTC
GTAGTAAGCAGGAACCTGTAAACAACACTGCAAGCAGGTAAAGGAGATGGGTGAATCACCACCTTGTAAGACATCCCCATGTTT
ATTTTGAAAACACACAAAACAATTATAACTATAGATAGATAAAATTGTTCTAAGTTTGATACATATATATTTTTTTTAATTTGAAA
AATTTTTTGATCTCGAGTCATTAATATGTAAATGCTAACAACGCTGATCTCGAAGTTAAAAGCAAATGTTTACTTTTCACATTAGA
CTAATTTTATAATACGCTCTGCAGTCTTCTATTTTGACATTTATATAAGTAAAATTAAATCATTTCAGAAACAATTTCTTTAGCGA
AATCCGAATATAGTTTTGCCGTCATATAAAAGCCAAGTAACACTTATTAAGCTTTCTAATCTTAAGAACAGAATTACTTCCACAA

ACATTTGTTTTTTAATAATCTGATATGCGATATTATTGGTTCTGTGGACCTTCAGATTAAGTTAGAATTTATTACTCCAATTAGTCC
TTTAAAGTGCAATATTTTCAGTCCAATTACGCAAATTTTTATATATTCCAAAGAATGTATACACGATTTGACATTTTAGTTTTAAT
AACTAGTATAGCAATTTTAAAAGTAGAATTATTGATTAAATATGTATTAAAGTTATATTTTATTTAAATGTGTATTGAAATAGTAT
TATTAATATAGCTTGTAAAAATTAATTTGCAAGCTTAACAATAATTTTAAAAATCAATAAATCCAAATATTTATACTTATATGGTT
ATAGTTATTTTATTAAAATAAAAATAAAATGTTAACTTATAAATTTGTTAGGATAGATAAGTGCAATTAATTTCAGAGTAGAGAT
TTATCAACCATAACAAAAATCTGCAATCAACATACCATATTTCAAAGACGATGGGAAATATCTACATACGTCAATCAACATACAA
TTTGGAGAGTGTTTTCTAAATATACATTAAATTCTACAATCAACATACCAACTACATGGCAATATTTCAATCAACTATAAAAATTA
TCAACAATCAACATTCCAAAATTGGACAATGATGATAAAATTTTAAAAATGTTACATCTGAGCAATATGTAATTTTTCTAAACAA
TAAAATTATGTATATAATTTATAAATGTCGAAATCAAATATAAATTAAAATTTTAAAAATATTGTTCCTCTAAAAAACGGTTACA
CTTGCCATAACGACTTGAATTAATATTGTAAAAAATAATTTAACTTCACAAAAATCTGTAAATTTATACTTTTCAAAATCTTGCCG
CAATGCAACATACTACTAAATTATATATTTTAAATTAAGAAATATTCATGGATATTTATCAACACTCAATGCCTAAGTCAAACTC
ATTTCTGTATATGATAAAATTTAATTGTTTTTCTCATGTTTAGGTGAAGGTTCTAGACCGTTATTTGAACGTGATGATATAACTGTC
AATCAGCCCACTGTAGGAGATTCACTCGAGAATATAACAAGCCATGATTGTAAAAATGGTTGTAAAGCTAAAAATTATAAAAGT
TCGTCAATTGTAAGTTGTGCCATTTAAGTGTTATAAGTACATAATAATGCTCTCTGTATATTAAGTTTATATTAATTTCCTTTGAAA
TATAAACAATTTGAAATTTAATTCATTTGTATAAAGTAGAAAATATGATCTAAAATTTACTTTTATACAAGTCAAATAGAAACAA
AAGATATTAATTGGATTATTTCAAAGGAATTTAATAAAACGTAAAACATGAGCTACAATTATTCTTACAAAACTTTTTTTAATATA
ACTGCAAAGGAATCGAGTTCACCTGAAAGATACCGACATCGAGATGCCAAAAAACGTTCTTCTCCACCCTCATCTGCGCGACGC
AAGTCACCTGAAAAATCTAGTCGTGATGAAAATTCGAGTCATAGTCATAATAGAAAGTACAATGAATATATAAAAAATAAGAGT
TGCAATACAAAGACGACTGAAAGATGTGGTAGAAGACGAGAAAAAAGCAAATCTCGTACACCATCGTCTACAAGAATAAAGAC
ATTAGCAGAGAAAAAACCATATTACAGGGATGAACAACGCGAAAGAGATCGTAGAAGACGTTTATATGGACATTCACGTTCTCG
TACACCTCCAATAATTTCAAGTAGAAGGTATAGCCCACACTTATCTAGAAGGCGTGTTGGCAGCCATACGAGAAGAAGGCGTTC
ATCTAGCAGAGATATAAATTATTCCCCTTATCATAGTTTCATGAGATATCATAGAAGTCGAAGATCTCGAACTCGTTCAAAAAGC

CGAACACCATCCTCAAGAAGAGACCGAGAACGTCACAAATCCCATTATAATAGAGAAAGACATAAAGAACAAAAGGATAATTC
AAATGCTTTAACAACTCCAGTTATTGCTGCTCCACCACCACAGATCATTCCAATAGCAGTTCCTGTACCGGCGGACTATGCTTATA
CATTTGTATGTTGCTAATTTATAGCAATAAATTTTTAGTATATAGCTTTTAATTTGTAAAATTATATCATTTTAGCCTGGTTGGACG
GCAACTCCGCAGGCATCCTGGCATACACCTCATCATCGGCCACCTACAGCTCCTACTCATTATGCTCCATTTCCTATGCTGCCAAT
GATGCCTCCACTTAGAGCTGCACATCGGACTCCATTTGGTGGTTTGCCGCCTCCATTACCTTATGGTCCAATGCCGCCTGGCTATC
GGCCGCACACTGGTCCAAGATATAATACCCCTAGGCATAATAACACACACAACTCGCATGCTCGTCCTAAAAAACCTACAAACT
AGAGAGGTGTTGTATGCATATTAATAAGCAATCAATGTATATATAATTAGTTTTGATTTAAAACTTTTGTATGGATGTTTAAAAAT
CTTTGTCGAAATGGAAGAGTGGCATATATTTATATGATATAAGACAAGTTACGAAAAAGGAACGTCTAAAACGTTTTCTTTTTTT
TCAATATACGTGAACTCTGTAATTTTTTTCCAGATTGTCTCTAAATATGATAAATCTGTGGTCATGTGATAGAATAATTTAAGCTA
CTCTTCCATATAGACACAGAAATAAAAAAATTTTGGCA

>Parent=Woma_00009986; type=transcript;
loc=ctg.000086F:complement(join(833397..833959,834030..834698, 834957..835093, 837412..837790)); name=female-specific *Wmtra* transcript; length=1748

GGCGTCAAGTTGATTAAGAGTGTAGCAAAATTTTCGGGAAGGTTAGGCTGTAGCTAAAAGTACAAATTGGATAAAATTGGAGAC
CTTGGCGACATTGTTGTTTTTATTACAAAACAAATAAATATTTAGTTTTTTGGGCCAAAAGGAAAGTTTCGCCTGTTACCATCAAC
ACCATCATCTATCGCCTTCTCAACTTTATTACGACCACAGCAACAATTAGAAATGGATTCCATAGCATCTGGCTTAGCAACATCA
GCTCTAAATATACCAAAGATTAAAATTCATCAATCTATACCAAGTGATTCCATAAGAAGAGGTCCCCGTGCCATCGTACGCTCAG
CTGCTTTAAACGACGACATTAATATTTTACGCAGATTTGGTGAAGGTTCTAGACCGTTATTTGAACGTGATGATATAACTGTCAAT
CAGCCCACTGTAGGAGATTCACTCGAGAATATAACAAGCCATGATTGTAAAAATGGTTGTAAAGCTAAAAATTATAAAAGTTCG
TCAATTGAATCGAGTTCACCTGAAAGATACCGACATCGAGATGCCAAAAAACGTTCTTCTCCACCCTCATCTGCGCGACGCAAGT
CACCTGAAAAATCTAGTCGTGATGAAAATTCGAGTCATAGTCATAATAGAAAGTACAATGAATATATAAAAAATAAGAGTTGCA
ATACAAAGACGACTGAAAGATGTGGTAGAAGACGAGAAAAAAGCAAATCTCGTACACCATCGTCTACAAGAATAAAGACATTA

GCAGAGAAAAAACCATATTACAGGGATGAACAACGCGAAAGAGATCGTAGAAGACGTTTATATGGACATTCACGTTCTCGTACA
CCTCCAATAATTTCAAGTAGAAGGTATAGCCCACACTTATCTAGAAGGCGTGTTGGCAGCCATACGAGAAGAAGGCGTTCATCTA
GCAGAGATATAAATTATTCCCCTTATCATAGTTTCATGAGATATCATAGAAGTCGAAGATCTCGAACTCGTTCAAAAAGCCGAAC
ACCATCCTCAAGAAGAGACCGAGAACGTCACAAATCCCATTATAATAGAGAAAGACATAAAGAACAAAAGGATAATTCAAATG
CTTTAACAACTCCAGTTATTGCTGCTCCACCACCACAGATCATTCCAATAGCAGTTCCTGTACCGGCGGACTATGCTTATACATTT
CCTGGTTGGACGGCAACTCCGCAGGCATCCTGGCATACACCTCATCATCGGCCACCTACAGCTCCTACTCATTATGCTCCATTTCC
TATGCTGCCAATGATGCCTCCACTTAGAGCTGCACATCGGACTCCATTTGGTGGTTTGCCGCCTCCATTACCTTATGGTCCAATGC
CGCCTGGCTATCGGCCGCACACTGGTCCAAGATATAATACCCCTAGGCATAATAACACACACAACTCGCATGCTCGTCCTAAAAA
ACCTACAAACTAGAGAGGTGTTGTATGCATATTAATAAGCAATCAATGTATATATAATTAGTTTTGATTTAAAACTTTTGTATGGA
TGTTTAAAAATCTTTGTCGAAATGGAAGAGTGGCATATATTTATATGATATAAGACAAGTTACGAAAAAGGAACGTCTAAAACG
TTTTCTTTTTTTTCAATATACGTGAACTCTGTAATTTTTTTCCAGATTGTCTCTAAATATGATAAATCTGTGGTCATGTGATAGAAT
AATTTAAGCTACTCTTCCATATAGACACAGAAATAAAAAAATTTTGGCA

>Parent=Woma_00009986; type=protein;
loc=ctg.000086F:complement(join(833690..833959,834030..834698, 834957..835093, 837412..837568)); name=female-specific WmTRA protein; length=410

MDSIASGLATSALNIPKIKIHQSIPSDSIRRGPRAIVRSAALNDDINILRRFGEGSRPLFERDDITVNQPTVGDSLENITSHDCKNGCKAKN
YKSSSIESSSPERYRHRDAKKRSSPPSSARRKSPEKSSRDENSSHSHNRKYNEYIKNKSCNTKTTERCGRRREKSKSRTPSSTRIKTLAEK
KPYYRDEQRERDRRRRLYGHSRSRTPPIISSRRYSPHLSRRRVGSHTRRRRSSSRDINYSPYHSFMRYHRSRRSRTRSKSRTPSSRRDRER
HKSHYNRERHKEQKDNSNALTTPVIAAPPPQIIPIAVPVPADYAYTFPGWTATPQASWHTPHHRPPTAPTHYAPFPMLPMMPPLRAAHR
TPFGGLPPPLPYGPMPPGYRPHTGPRYNTPRHNNTHNSHARPKKPTN

>Parent=Woma_00009986; type= transcript;
loc=ctg.000086F:complement(join(833397..833959,834030..834698, 834957..835093, 837134..837790)); name=male-specific *Wmtra* transcript;

length=2026

GGCGTCAAGTTGATTAAGAGTGTAGCAAAATTTTCGGGAAGGTTAGGCTGTAGCTAAAAGTACAAATTGGATAAAATTGGAGAC
CTTGGCGACATTGTTGTTTTTATTACAAAACAAATAAATATTTAGTTTTTTGGGCCAAAAGGAAAGTTTCGCCTGTTACCATCAAC
ACCATCATCTATCGCCTTCTCAACTTTATTACGACCACAGCAACAATTAGAAATGGATTCCATAGCATCTGGCTTAGCAACATCA
GCTCTAAATATACCAAAGATTAAAATTCATCAATCTATACCAAGTGATTCCATAAGAAGAGGTCCCCGTGCCATCGTACGCTCAG
CTGCTTTAAACGACGACATTAATATTTTACGCAGATTTGGTAATTTTCTTACAAAATTTCTGCATACCTTTTAAATTTTTTTGCAGT
ATTTAATACAATGAGAGTAGTTACAAAGAAATTAACAACATTATAAGGAACGTAAAAGAAATTAGATATTTTTATCAGTAATTAT
TGAATGGCACAATGAAAGAAAAAAGCTACTATTGAAGGCTGTGACGCGATATGCAGCGCTAGAAAAATTCATCAATCAACATTC
CATATCCATTTCCTTGAACAGTGAGGAGAATCTTCAAACAACTAAACCGCTGACTACAACTGTGAAGGTTCTAGACCGTTATTTG
AACGTGATGATATAACTGTCAATCAGCCCACTGTAGGAGATTCACTCGAGAATATAACAAGCCATGATTGTAAAAATGGTTGTA
AAGCTAAAAATTATAAAAGTTCGTCAATTGAATCGAGTTCACCTGAAAGATACCGACATCGAGATGCCAAAAAACGTTCTTCTCC
ACCCTCATCTGCGCGACGCAAGTCACCTGAAAAATCTAGTCGTGATGAAAATTCGAGTCATAGTCATAATAGAAAGTACAATGA
ATATATAAAAAATAAGAGTTGCAATACAAAGACGACTGAAAGATGTGGTAGAAGACGAGAAAAAAGCAAATCTCGTACACCAT
CGTCTACAAGAATAAAGACATTAGCAGAGAAAAAACCATATTACAGGGATGAACAACGCGAAAGAGATCGTAGAAGACGTTTA
TATGGACATTCACGTTCTCGTACACCTCCAATAATTTCAAGTAGAAGGTATAGCCCACACTTATCTAGAAGGCGTGTTGGCAGCC
ATACGAGAAGAAGGCGTTCATCTAGCAGAGATATAAATTATTCCCCTTATCATAGTTTCATGAGATATCATAGAAGTCGAAGATC
TCGAACTCGTTCAAAAAGCCGAACACCATCCTCAAGAAGAGACCGAGAACGTCACAAATCCCATTATAATAGAGAAAGACATAA
AGAACAAAAGGATAATTCAAATGCTTTAACAACTCCAGTTATTGCTGCTCCACCACCACAGATCATTCCAATAGCAGTTCCTGTA
CCGGCGGACTATGCTTATACATTTCCTGGTTGGACGGCAACTCCGCAGGCATCCTGGCATACACCTCATCATCGGCCACCTACAG
CTCCTACTCATTATGCTCCATTTCCTATGCTGCCAATGATGCCTCCACTTAGAGCTGCACATCGGACTCCATTTGGTGGTTTGCCG
CCTCCATTACCTTATGGTCCAATGCCGCCTGGCTATCGGCCGCACACTGGTCCAAGATATAATACCCCTAGGCATAATAACACAC
ACAACTCGCATGCTCGTCCTAAAAAACCTACAAACTAGAGAGGTGTTGTATGCATATTAATAAGCAATCAATGTATATATAATTA

GTTTTGATTTAAAACTTTTGTATGGATGTTTAAAAATCTTTGTCGAAATGGAAGAGTGGCATATATTTATATGATATAAGACAAGT
TACGAAAAAGGAACGTCTAAAACGTTTTCTTTTTTTTCAATATACGTGAACTCTGTAATTTTTTTCCAGATTGTCTCTAAATATGAT
AAATCTGTGGTCATGTGATAGAATAATTTAAGCTACTCTTCCATATAGACACAGAAATAAAAAAATTTTGGCA
>Parent=Woma_00009986; type=protein; loc=ctg.000086F:complement(837377..837568); name=male-specific WmTRA protein; length=63
MDSIASGLATSALNIPKIKIHQSIPSDSIRRGPRAIVRSAALNDDINILRRFGNFLTKFLHTF

File S2. Gene, transcript, and protein sequences of *Wmtra2*.

> ID=Woma_00008264; type=gene; loc=ctg.000059F:3139052..3142296; name=*Wmtra2*; length=3116

ATAGACGGCTTTGCTTTTGTACAGTGTTTTGATAAATTTCTATTTTGTTTAAATCAAATATTTTTTTTGGCGAAATTGAACATTTAC
GGAGGAAAAATTAAGATTTATTATAATTTAAACAAACTAAGCTTTATATTAAGAGTTACAGAACAAAATGGTAAATTATTTCCCC
TGGATACTTTCTCATCAAAAATCATTTGATTAATTATAATTTAAAGCATTTATGTATTCAATTTACATATTTTTCTTGTACTATTTTC
AGAGTCCTCGTTCACGCAGTCGTTCCATAACACCACGTCGTTCATATTCTCGCTCACCCTATAGACGCAGCGGACGCCGGTAAGT
AGTGTAAAATTGCCCTTTAGTTATTGGTTATATTTGATTACAATTTTCTTCTCTGATTGTATAAATGAAGAAAATCTTACAGCCGT
CATAAGTATGACTCTCGTAGTAGTTCCAGTTCGCATCCGCCGTCGCCACCACCACTTACATCTAGGCGACACAATAGCCGTTATTC
CAGCCGCAGTATATCAAGGTAAGTTCTGCTTAATTTGCATCAAATTTAGAGGTTAGGTTGAGCACATTTGAAAATTTTAAAAAAA
TTTTTATTAATTGAATAGCTTGGATAAAATAATTCTAAATGTGTATATATATATATATATATATATATATATATATATATATATA
TAAAAAACATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT
ATATATATATATATATATATATGTTTTTTGAAAACTGCCTTTTTATGGTTTTCTTAAAGTAGATATAAATAATGGATGGATTTTTAACG
CCTGTCTATTTTTAAAACTTAAAATATTTTAAGTTTGGTTAAATAATGCAGGTACTTTTATAGCTTAAGAATTAATTTCAGAGAAA
GCTAATTTAACTCGTATTTTCTTTATTATTTAGTTCAGGGCTTAAAATAATTGTTTCATTCAAAATATTTCATAAAGTTATCATAAT
ATTGTATATAGTATGACTCATTTTATTATATAAAATTGTAACATATTGATAATTTATTTTTAGGAATTTGAATTCTGAATTTTTTCT
ATTCAAATCAAACTATTGAAATAATTAAAGTTTTTAGCAAATTTCAACTAAACAACATGCAGCAATTAAAATGGCAAACTCAAAA
TAATATATAGAAACGATGCAAGTAACTTCAAATACTTTCAGGTTTTCTTTGTTTTTCAAAAACTAATCTAAAATATTGAAAAATCA
AAGTGTAGCAAATTTCCATTACAAAACATAACAATTATGGTACAAGAAAACCATACAACTTGGCAAAATTCTTTGAAAAACTAA
GGATTTTTGTCAAATCAATGAAACTTAATTTGGAAATAATTCGAAGAATAATGCAAAACAAAAAACGAACATATTGAAATAAAC
TTGAATATAAATCGTATTCGAAATACTAACGTTAATAATTTCTTTTTCGTTTTGTGTAATATTAAAATTTTTATTATTTTAATCCTG
CAGATCCATATCACCCGATTATAGAAAACGTTATACTAGTCGTAGGCATTATTCGAAATCTCGTTCGAAAACCAAAAGCCGTTCA
AGATCACGTTCACGCACATCGTACGATAAAAATATACGCAATAATAGAGTAAGTAAATTGGAATATACATTTGAATTTTTATAAT

TTAACCAATCTAAAAAATTGAATTGAAATTGAGAGTTGAAGCAGTTTGTAAAAGTTTAGTTTAATTATATACTTTTAGGAAAAAC
CTTTGCCTTGCCGTTGTATTGGAGTATTTGGTTTAAGCGTTTATACTACACAACAAAAAATAAGAGAAATATTTTCAAAATTTGGT
CCTATCGAACGTATACAAGTAATAACTGATGCGCAGGTAAGGTTTATTTGCAATATTTTTGTCCATTTAATTAAAAGCCTTTTATT
TTATTCATTTTATTTAAAGTTGTAAGAAATAAAATGACACTCGAAGAGGCTAAAATAATTATTTTGAACAATAATAAATAGAGTT
TTGATTTAAGAATATTATAAACAAAACCAAAGGCTTGTTGTTACACTATAAACGTAGTAACGTCGTTTTTTCAATACTGTTTTTAT
AGTTTTTTGTACTGTTTTTTTATTTTATTTTTTTCATAGACCCGTCGTTCGCGTGGTTTTTGTTTTATTTATTATGATAATTTAGCTGA
TGCTAGAGCCGCCTGTGATAACTGTTGTGGCATGGAAATTGATGGCAGACGTATACGTGTTGCCTATTCCATAACTGAAAGGCCA
CACACACCCACTCCTGGTGTTTATATGGGAAAACCAGCAAAGTAAATATTCAATATTGAAATTTTTACATTTAAAATATTGTTTTG
TTGCATTTTTAGAGATTTAAGAGAACGTTATCGTGCCCAAAAGCAACAGCGCCATCATTCTCCAATTACTTATAAAAGTCACCGT
AGTAGTAACCGCCATCATTATGAACGCAGTCGTAGCCGTTCCTATACACCACGTAAGTAATGTTTTCATTATTTTCTATTACAATT
TCTTATACCTCAGGAGCGTTGTTAAGTTGTTGCGCCATTTTAATTTTAACTCCTTATTTATATGCTTTATGCTTTTGCTGCTACATTT
TTGATTTGTAAATCATTTCAAAACATAGAATTGAGTGTAAAGAATGAATTCAAAATATTAATTTTTAATATCTTTTACAGGTCGTT
ATCGTTACTGATGCCCCTTGAAGTGTGTCAGTATTTATTCCATCGAACCATATGCATAATTTTTGAATGCTCTTAATTTTTGAATAG
TTTCTCTTTTTTACTTTTGCTTGTGTATTTTCATCATTTCATTTGTTTGTTTTTTTTTTTATTATTTATTGGTATCGAGATTGAATTTTTT
TCTGTTTTATGATTTATTGTTTAGTTTTAATTGTAACATAGAATATATTGAAAATATCATTTTTTCTGTTTGAATACAGATTTCCTTA
ATTTTCTGGTAAATTTTTGTTTTATAAAAATTAATGAAATACAACAAAAAAGAGAAGCATTTCGTTTAAGAACGAATGCTTCTATT
TATTTCCTTAACAT

>Parent=Woma_00008264; type= transcripts;
loc=ctg.000059F:join(3139052..3139208,3139314..3139390,3139467..3139586,3140603..3140732,3140848..3140976,3141237..3141412,31414
70..3141594,3141797..3142167); name=*Wmtra2* transcripts; length=1285

ATAGACGGCTTTGCTTTTGTACAGTGTTTTGATAAATTTCTATTTTGTTTAAATCAAATATTTTTTTTGGCGAAATTGAACATTTAC
GGAGGAAAAATTAAGATTTATTATAATTTAAACAAACTAAGCTTTATATTAAGAGTTACAGAACAAAATGAGTCCTCGTTCACGC

AGTCGTTCCATAACACCACGTCGTTCATATTCTCGCTCACCCTATAGACGCAGCGGACGCCGAAAATCTTACAGCCGTCATAAGT
ATGACTCTCGTAGTAGTTCCAGTTCGCATCCGCCGTCGCCACCACCACTTACATCTAGGCGACACAATAGCCGTTATTCCAGCCG
CAGTATATCAAGATCCATATCACCCGATTATAGAAAACGTTATACTAGTCGTAGGCATTATTCGAAATCTCGTTCGAAAACCAAA
AGCCGTTCAAGATCACGTTCACGCACATCGTACGATAAAAATATACGCAATAATAGAGAAAAACCTTTGCCTTGCCGTTGTATTG
GAGTATTTGGTTTAAGCGTTTATACTACACAACAAAAAATAAGAGAAATATTTTCAAAATTTGGTCCTATCGAACGTATACAAGT
AATAACTGATGCGCAGACCCGTCGTTCGCGTGGTTTTTGTTTTATTTATTATGATAATTTAGCTGATGCTAGAGCCGCCTGTGATA
ACTGTTGTGGCATGGAAATTGATGGCAGACGTATACGTGTTGCCTATTCCATAACTGAAAGGCCACACACACCCACTCCTGGTGT
TTATATGGGAAAACCAGCAAAAGATTTAAGAGAACGTTATCGTGCCCAAAAGCAACAGCGCCATCATTCTCCAATTACTTATAA
AAGTCACCGTAGTAGTAACCGCCATCATTATGAACGCAGTCGTAGCCGTTCCTATACACCACGTCGTTATCGTTACTGATGCCCCT
TGAAGTGTGTCAGTATTTATTCCATCGAACCATATGCATAATTTTTGAATGCTCTTAATTTTTGAATAGTTTCTCTTTTTTACTTTTG
CTTGTGTATTTTCATCATTTCATTTGTTTGTTTTTTTTTTATTATTTATTGGTATCGAGATTGAATTTTTTTCTGTTTTATGATTTATT
GTTTAGTTTTAATTGTAACATAGAATATATTGAAAATATCATTTTTTCTGTTTGAATACAGATTTCCTTAATTTTCTGGTAAATTTT
TGTTTTATAAAAATTAATGAAATACAACAAAAAAGAGAAGCATTTCGTTTAAGAACGAATGCTTCTATTTATTTCCTTAACAT
>Parent= Woma_00008264; type=protein;
loc=ctg.000059F:join(3139206..3139208,3139314..3139390,3139467..3139586,3140603..3140732,3140848..3140976,3141237..3141412,31414
70..3141594,3141797.. 3141813); name=WmTRA2 protein; length=258
MSPRSRSRSITPRRSYSRSPYRRSGRRKSYSRHKYDSRSSSSSHPPSPPPLTSRRHNSRYSSRSISRSISPDYRKRYTSRRHYSKSRSKTKSR
SRSRSRTSYDKNIRNNREKPLPCRCIGVFGLSVYTTQQKIREIFSKFGPIERIQVITDAQTRRSRGFCFIYYDNLADARAACDNCCGMEID
GRRIRVAYSITERPHTPTPGVYMGKPAKDLRERYRAQKQQRHHSPITYKSHRSSNRHHYERSRSRSYTPRRYRY

File S3. U6 RNA sequences in *W. magnifica*.

> U6 spliceosomal RNA; Loc=ctg.000025F: 1221473..1221568

GTTCTTGCTTCGGCAGAACATATACTAAAATGGGAACGATACAGAGAAGATTAGCATGGCCCCTACACGCAAAATCGTGAAGCG
TTCCACATTTTT

> U6 spliceosomal RNA; Loc=ctg.000033F: 5094703..5094813

GTTCTTGCTTCGGTAGAACATATATATACTAAAATTGGAACGATACAGAGAAGATTAGCATGGCCTTTGGGCACGGGTGACACGC
AAAATCGTAAAGCCTTCCACTTTTTT

> U6 spliceosomal RNA; Loc=ctg.000034F: 542616..542719

AATGTCAATCCGACCTCGTACACTAAAATTGGAACGATACAGAGAAGATTAGCATGGCCCTGCGTAAGTATGACTCGCAAAATC
GTGAAGTGTGCCACATGTTT

> U6 spliceosomal RNA; Loc=ctg.000034F: 2328939..2329037

GCTCTTGCTTCGGCAGAACATATACTAAAATTGGAACGATACAGAGAATATTAGCATGGCTCTGCGTAAGGATGACACGTAAAA
TCGTTCCACATTTTT

>U6 spliceosomal RNA; Loc=ctg.000037F: 2976518..2976613 reverse complement

GTTTTTCTGGCTTCGGTAGAACATACACTAAAATTGGAAGATACAGAGAAGATTAGCAAGGATGACACGCAAAATTTTGAAGCG
TTCCACATTTTT

>U6 spliceosomal RNA; Loc=ctg.000437F: 403325..403430 reverse complement

CTTCTTGCTTCGGCAGAACATATACTAAAATTGGAACGATACAGAGAAGATCAGCATGGCCCTGTGCAAGGATGACACGCAAAA
TCGTGAAGCGTTCCACATTTTT

Table S1. Statistics of RNA-seq data for adult females and adult males.

Samples	Raw bases (G)	Q20 (%)	Q30 (%)	GC Content (%)	Clean bases (G)	Clean reads	Genome mapping rate (%)
Female_1	6.66	97.7	93.4	37.2	6.44	43,325,756	94.77
Female_2	6.39	97.7	93.4	39.5	6.17	41,603,624	95.73
Female_3	6	97.8	93.7	34.3	5.83	39,186,716	94.75
Male_1	7.28	97.6	93.1	35.4	6.98	47,109,424	93.05
Male_2	6.74	97.6	92.9	35.3	6.48	43,736,892	92.35
Male_3	7.02	97.3	92.5	36.4	6.76	45,586,092	94.06

Table S4. Expression levels of five female-biased genes.

Gene_id	Woma_00004829	Woma_00005371	Woma_00011211	Woma_00012414	Woma_00010258
Gene_name	*vas*	*nanos*	*bcd*	*BicC*	*inx5*
Count_Female_1	901	90	20	62	199
Count_Female_2	4210	315	62	571	995
Count_Female_3	571	101	6	38	845
Count_Male_1	85	0	0	0	20
Count_Male_2	83	0	0	0	0
Count_Male_3	130	1	0	0	0
TPM_Female_1	11.39	4.22	0.95	1.33	10.21
TPM_Female_2	56.42	15.66	3.14	13.01	54.14
TPM_Female_3	7.15	4.69	0.28	0.81	42.95
TPM_Male_1	1.36	0	0	0	1.3
TPM_Male_2	1.22	0	0	0	0
TPM_Male_3	1.66	0.05	0	0	0

Table S5. *W. magnifica* orthologs of essential genes in *D. melanogaster*.

Wohlfahrtia magnifica Gene id	*Drosophila melanogaster* Gene id	Phenotype
Woma_00013373	FBgn0037757	lethal
Woma_00010373	FBgn0039481	lethal
Woma_00000425	FBgn0000038	partially lethal
Woma_00009408	FBgn0086656	lethal
Woma_00000356	FBgn0263108	partially lethal
Woma_00008309	FBgn0004511	lethal
Woma_00000998	FBgn0033787	lethal
Woma_00003193	FBgn0052069	lethal
Woma_00001389	FBgn0000606	lethal
Woma_00003551	FBgn0037099	lethal
Woma_00012283	FBgn0003460	lethal
Woma_00013843	FBgn0284251	partially lethal
Woma_00011839	FBgn0011211	lethal
Woma_00001775	FBgn0037835	partially lethal
Woma_00010160	FBgn0010282	lethal
Woma_00012731	FBgn0030061	lethal
Woma_00015140	FBgn0035947	lethal
Woma_00010004	FBgn0002526	lethal
Woma_00000612	FBgn0039385	lethal
Woma_00007095	FBgn0036334	partially lethal
Woma_00000476	FBgn0035372	lethal
Woma_00012344	FBgn0011739	lethal
Woma_00006815	FBgn0026582	lethal
Woma_00010961	FBgn0003978	lethal
Woma_00009807	FBgn0039562	lethal
Woma_00002095	FBgn0029840	lethal
Woma_00006700	FBgn0000377	lethal
Woma_00012221	FBgn0034231	lethal
Woma_00010216	FBgn0035769	partially lethal
Woma_00004973	FBgn0051111	lethal
Woma_00003953	FBgn0040534	lethal
Woma_00012555	FBgn0050349	partially lethal
Woma_00007032	FBgn0035620	partially lethal
Woma_00011945	FBgn0283494	lethal
Woma_00005031	FBgn0284255	lethal
Woma_00012618	FBgn0032691	lethal
Woma_00005785	FBgn0024352	lethal
Woma_00007942	FBgn0028690	lethal
Woma_00012131	FBgn0034940	lethal
Woma_00012299	FBgn0025830	lethal
Woma_00002841	FBgn0037342	lethal
Woma_00014501	FBgn0032138	lethal
Woma_00002458	FBgn0004170	lethal
Woma_00010321	FBgn0003655	lethal
Woma_00002473	FBgn0011758	lethal
Woma_00010202	FBgn0040298	lethal
Woma_00010593	FBgn0031886	lethal
Woma_00005699	FBgn0004110	lethal
Woma_00004026	FBgn0031488	lethal
Woma_00012093	FBgn0028744	lethal
Woma_00005429	FBgn0023175	lethal
Woma_00010518	FBgn0023174	lethal
Woma_00003511	FBgn0037345	partially lethal
Woma_00009920	FBgn0030670	lethal
Woma_00007378	FBgn0005355	lethal

Woma_00002681	FBgn0020272	lethal
Woma_00011353	FBgn0034401	lethal
Woma_00003284	FBgn0035443	partially lethal
Woma_00011554	FBgn0010803	lethal
Woma_00013734	FBgn0026722	lethal
Woma_00014484	FBgn0034346	partially lethal
Woma_00002817	FBgn0010774	lethal
Woma_00000281	FBgn0015805	lethal
Woma_00001743	FBgn0037664	partially lethal
Woma_00012064	FBgn0004132	lethal
Woma_00004932	FBgn0032444	lethal
Woma_00005695	FBgn0011278	lethal
Woma_00002635	FBgn0028693	lethal
Woma_00008944	FBgn0036761	lethal
Woma_00011208	FBgn0000439	lethal
Woma_00006080	FBgn0015618	lethal
Woma_00009071	FBgn0027500	lethal
Woma_00009167	FBgn0001986	lethal
Woma_00006166	FBgn0034908	lethal
Woma_00012330	FBgn0039561	partially lethal
Woma_00006994	FBgn0036829	partially lethal
Woma_00004015	FBgn0266670	lethal
Woma_00010117	FBgn0024222	partially lethal
Woma_00007465	FBgn0261479	lethal
Woma_00011146	FBgn0004650	lethal
Woma_00001327	FBgn0086712	lethal
Woma_00003007	FBgn0028970	lethal
Woma_00002420	FBgn0030305	lethal
Woma_00004601	FBgn0033378	lethal
Woma_00008031	FBgn0004378	lethal
Woma_00006677	FBgn0028836	partially lethal
Woma_00004167	FBgn0001995	lethal
Woma_00002457	FBgn0002561	lethal
Woma_00014406	FBgn0033272	lethal
Woma_00010396	FBgn0040283	lethal
Woma_00007489	FBgn0039013	partially lethal
Woma_00012690	FBgn0045035	lethal
Woma_00012864	FBgn0037229	lethal
Woma_00008518	FBgn0031822	lethal
Woma_00014424	FBgn0032409	lethal
Woma_00008000	FBgn0044028	lethal
Woma_00011930	FBgn0283678	partially lethal
Woma_00005949	FBgn0000618	partially lethal
Woma_00006548	FBgn0045502	partially lethal
Woma_00008456	FBgn0001967	lethal
Woma_00007747	FBgn0032924	lethal
Woma_00015343	FBgn0000533	lethal
Woma_00007950	FBgn0027497	lethal
Woma_00005532	FBgn0053110	lethal
Woma_00003937	FBgn0025825	lethal
Woma_00014842	FBgn0010808	lethal
Woma_00000764	FBgn0038649	partially lethal
Woma_00013325	FBgn0019644	partially lethal
Woma_00003536	FBgn0261801	partially lethal
Woma_00001390	FBgn0034061	partially lethal
Woma_00009603	FBgn0028484	lethal
Woma_00005524	FBgn0040078	lethal
Woma_00010623	FBgn0040079	partially lethal

Woma_00005449	FBgn0266723	partially lethal
Woma_00000816	FBgn0039639	lethal
Woma_00014192	FBgn0027535	lethal
Woma_00002828	FBgn0261049	lethal
Woma_00013955	FBgn0052528	lethal
Woma_00005322	FBgn0039132	lethal
Woma_00001070	FBgn0033032	lethal
Woma_00010129	FBgn0051344	lethal
Woma_00016402	FBgn0036570	partially lethal
Woma_00007532	FBgn0261596	lethal
Woma_00004419	FBgn0039125	lethal
Woma_00009973	FBgn0260859	lethal
Woma_00010943	FBgn0262117	lethal
Woma_00005759	FBgn0036811	lethal
Woma_00004395	FBgn0039691	lethal
Woma_00003484	FBgn0038585	lethal
Woma_00011122	FBgn0035956	lethal
Woma_00003928	FBgn0030864	partially lethal
Woma_00013441	FBgn0011785	lethal
Woma_00004885	FBgn0263933	lethal
Woma_00013462	FBgn0039403	lethal
Woma_00006362	FBgn0028999	lethal
Woma_00004487	FBgn0033753	partially lethal
Woma_00003735	FBgn0024491	lethal
Woma_00012741	FBgn0004864	lethal
Woma_00004194	FBgn0031454	lethal
Woma_00008332	FBgn0020909	partially lethal
Woma_00014412	FBgn0015544	lethal
Woma_00001807	FBgn0023212	lethal
Woma_00015334	FBgn0028684	lethal
Woma_00005698	FBgn0004862	lethal
Woma_00002035	FBgn0014868	lethal
Woma_00009059	FBgn0025582	lethal
Woma_00003203	FBgn0035997	lethal
Woma_00006117	FBgn0042092	partially lethal
Woma_00006164	FBgn0010278	lethal
Woma_00000455	FBgn0003206	lethal
Woma_00000446	FBgn0035851	lethal
Woma_00002689	FBgn0037561	lethal
Woma_00003346	FBgn0037364	partially lethal
Woma_00003855	FBgn0038768	lethal
Woma_00011523	FBgn0020388	lethal
Woma_00013543	FBgn0028737	lethal
Woma_00012883	FBgn0039620	lethal
Woma_00002593	FBgn0260771	partially lethal
Woma_00012651	FBgn0033731	lethal
Woma_00007514	FBgn0010768	lethal
Woma_00003899	FBgn0002609	lethal
Woma_00006992	FBgn0036839	partially lethal
Woma_00000941	FBgn0016641	lethal
Woma_00009503	FBgn0021944	lethal
Woma_00011170	FBgn0029848	lethal
Woma_00013310	FBgn0039139	partially lethal
Woma_00006990	FBgn0052442	lethal
Woma_00007764	FBgn0000482	lethal
Woma_00011515	FBgn0035526	partially lethal
Woma_00001226	FBgn0026261	lethal
Woma_00015844	FBgn0032228	lethal

Woma_00006203	FBgn0034488	partially lethal
Woma_00005288	FBgn0051133	lethal
Woma_00003908	FBgn0003312	lethal
Woma_00015346	FBgn0013269	partially lethal
Woma_00016626	FBgn0032704	partially lethal
Woma_00010703	FBgn0031260	lethal
Woma_00010879	FBgn0004407	lethal
Woma_00000959	FBgn0033629	partially lethal
Woma_00013370	FBgn0028374	lethal
Woma_00002034	FBgn0040235	partially lethal
Woma_00008179	FBgn0259231	lethal
Woma_00005515	FBgn0028688	lethal
Woma_00011669	FBgn0033539	partially lethal
Woma_00012553	FBgn0002552	lethal
Woma_00011444	FBgn0028427	lethal
Woma_00003573	FBgn0040347	lethal
Woma_00014898	FBgn0032260	lethal
Woma_00005677	FBgn0026634	lethal
Woma_00010380	FBgn0039486	lethal
Woma_00014258	FBgn0022097	lethal
Woma_00007089	FBgn0036386	partially lethal
Woma_00011150	FBgn0024364	lethal
Woma_00014305	FBgn0030035	lethal
Woma_00004262	FBgn0039754	lethal
Woma_00000457	FBgn0035499	partially lethal
Woma_00011331	FBgn0026602	partially lethal
Woma_00010817	FBgn0260985	lethal
Woma_00006291	FBgn0034727	lethal
Woma_00000440	FBgn0010223	partially lethal
Woma_00008927	FBgn0036772	lethal
Woma_00010935	FBgn0028538	lethal
Woma_00001197	FBgn0013548	lethal
Woma_00014014	FBgn0031209	partially lethal
Woma_00007193	FBgn0033657	lethal
Woma_00006455	FBgn0024196	lethal
Woma_00004801	FBgn0264810	lethal
Woma_00004360	FBgn0037684	partially lethal
Woma_00000867	FBgn0005642	lethal
Woma_00010751	FBgn0003448	lethal
Woma_00008445	FBgn0260011	partially lethal
Woma_00004614	FBgn0028506	partially lethal
Woma_00000115	FBgn0025644	partially lethal
Woma_00007026	FBgn0036206	partially lethal
Woma_00014463	FBgn0033264	lethal
Woma_00011009	FBgn0033557	lethal
Woma_00012738	FBgn0030851	lethal
Woma_00005479	FBgn0013717	lethal
Woma_00001818	FBgn0024509	lethal
Woma_00005958	FBgn0038597	lethal
Woma_00002903	FBgn0045980	lethal
Woma_00009983	FBgn0035692	lethal
Woma_00005626	FBgn0037534	lethal
Woma_00011184	FBgn0023521	lethal
Woma_00012845	FBgn0015331	lethal
Woma_00006154	FBgn0022800	lethal
Woma_00003600	FBgn0024194	lethal
Woma_00004024	FBgn0031492	lethal
Woma_00008559	FBgn0261445	partially lethal

Woma_00006520	FBgn0011236	lethal
Woma_00002239	FBgn0037466	lethal
Woma_00007326	FBgn0037573	lethal
Woma_00003684	FBgn0020392	lethal
Woma_00003713	FBgn0263832	lethal
Woma_00010014	FBgn0034817	lethal
Woma_00002934	FBgn0032988	lethal
Woma_00002263	FBgn0250874	lethal
Woma_00015623	FBgn0020270	lethal
Woma_00012127	FBgn0015621	lethal
Woma_00004489	FBgn0086908	lethal
Woma_00001808	FBgn0038806	lethal
Woma_00007659	FBgn0024188	lethal
Woma_00006390	FBgn0026252	lethal
Woma_00006605	FBgn0014877	lethal
Woma_00002423	FBgn0030286	partially lethal
Woma_00006355	FBgn0036643	lethal
Woma_00005387	FBgn0039175	lethal
Woma_00003409	FBgn0038380	partially lethal
Woma_00003661	FBgn0036274	partially lethal
Woma_00011593	FBgn0005533	partially lethal
Woma_00003792	FBgn0037632	lethal
Woma_00004016	FBgn0031536	lethal
Woma_00006386	FBgn0035762	lethal
Woma_00009866	FBgn0040385	partially lethal
Woma_00010954	FBgn0003141	lethal
Woma_00003895	FBgn0002733	lethal
Woma_00013160	FBgn0040348	partially lethal
Woma_00010883	FBgn0001965	lethal
Woma_00004634	FBgn0031601	lethal
Woma_00002786	FBgn0051390	lethal
Woma_00002830	FBgn0261286	lethal
Woma_00003896	FBgn0002732	lethal
Woma_00007689	FBgn0032200	partially lethal
Woma_00009690	FBgn0025626	lethal
Woma_00013362	FBgn0024293	partially lethal
Woma_00004332	FBgn0038927	lethal
Woma_00005497	FBgn0002565	lethal
Woma_00013393	FBgn0031661	lethal
Woma_00006123	FBgn0026196	lethal
Woma_00006439	FBgn0036264	lethal
Woma_00006459	FBgn0034585	lethal
Woma_00008188	FBgn0001079	lethal
Woma_00003957	FBgn0026380	lethal
Woma_00006516	FBgn0025352	lethal
Woma_00004118	FBgn0262467	lethal
Woma_00003432	FBgn0038016	lethal
Woma_00009711	FBgn0025463	lethal
Woma_00005471	FBgn0039030	lethal
Woma_00009723	FBgn0266282	lethal
Woma_00003832	FBgn0016691	lethal
Woma_00001772	FBgn0017577	lethal
Woma_00003960	FBgn0260243	lethal
Woma_00008583	FBgn0035148	partially lethal
Woma_00003250	FBgn0027055	partially lethal
Woma_00005247	FBgn0029088	lethal
Woma_00002793	FBgn0037637	lethal
Woma_00012262	FBgn0034142	lethal

Woma_00000157	FBgn0030086	lethal
Woma_00004944	FBgn0266369	lethal
Woma_00006581	FBgn0027589	partially lethal
Woma_00012916	FBgn0263005	lethal
Woma_00010735	FBgn0261020	lethal
Woma_00004383	FBgn0038808	lethal
Woma_00004382	FBgn0039016	lethal
Woma_00009068	FBgn0004556	lethal
Woma_00011526	FBgn0036298	lethal
Woma_00003571	FBgn0036698	lethal
Woma_00006881	FBgn0036775	lethal
Woma_00008803	FBgn0040346	lethal
Woma_00006985	FBgn0026371	lethal
Woma_00003741	FBgn0003205	lethal
Woma_00010469	FBgn0019938	partially lethal
Woma_00011553	FBgn0037244	lethal
Woma_00000459	FBgn0036038	lethal
Woma_00007220	FBgn0001321	lethal
Woma_00001010	FBgn0033755	lethal
Woma_00011887	FBgn0031801	partially lethal
Woma_00012074	FBgn0025373	lethal
Woma_00003799	FBgn0024833	lethal
Woma_00000812	FBgn0039731	partially lethal
Woma_00005408	FBgn0034494	lethal
Woma_00004210	FBgn0031573	partially lethal
Woma_00015681	FBgn0031399	partially lethal
Woma_00014959	FBgn0021995	lethal
Woma_00003125	FBgn0263038	lethal
Woma_00009460	FBgn0032194	lethal
Woma_00001801	FBgn0038126	lethal
Woma_00001641	FBgn0033769	partially lethal
Woma_00000131	FBgn0263237	partially lethal
Woma_00005803	FBgn0028694	lethal
Woma_00001855	FBgn0039258	lethal
Woma_00000258	FBgn0029913	partially lethal
Woma_00007951	FBgn0037371	lethal
Woma_00006174	FBgn0264962	lethal
Woma_00013319	FBgn0002775	lethal
Woma_00009360	FBgn0042083	lethal
Woma_00013102	FBgn0004106	lethal
Woma_00009979	FBgn0035965	lethal
Woma_00005952	FBgn0027090	lethal
Woma_00006934	FBgn0036813	lethal
Woma_00009848	FBgn0038269	lethal
Woma_00012259	FBgn0033556	lethal
Woma_00005243	FBgn0038903	partially lethal
Woma_00012903	FBgn0031086	lethal
Woma_00006528	FBgn0029175	lethal
Woma_00002649	FBgn0260858	lethal
Woma_00011525	FBgn0004926	lethal
Woma_00009575	FBgn0000206	lethal
Woma_00005760	FBgn0036812	lethal
Woma_00008138	FBgn0053505	lethal
Woma_00003857	FBgn0263863	lethal
Woma_00012835	FBgn0015400	partially lethal
Woma_00000495	FBgn0033842	lethal
Woma_00002086	FBgn0023509	partially lethal
Woma_00003961	FBgn0039337	lethal

Woma_00015261	FBgn0026873	partially lethal
Woma_00008134	FBgn0025393	lethal
Woma_00014475	FBgn0261119	lethal
Woma_00008589	FBgn0266724	lethal
Woma_00013631	FBgn0032343	partially lethal
Woma_00012607	FBgn0032059	partially lethal
Woma_00010238	FBgn0035785	partially lethal
Woma_00006226	FBgn0050476	lethal
Woma_00008061	FBgn0035141	lethal
Woma_00012165	FBgn0038002	partially lethal
Woma_00005988	FBgn0266674	lethal
Woma_00013053	FBgn0033129	partially lethal
Woma_00002105	FBgn0000524	lethal
Woma_00012596	FBgn0020305	lethal
Woma_00009491	FBgn0051876	lethal
Woma_00014816	FBgn0261016	partially lethal
Woma_00010454	FBgn0267824	lethal
Woma_00012324	FBgn0039558	lethal
Woma_00009673	FBgn0037549	partially lethal
Woma_00007425	FBgn0086695	lethal
Woma_00012500	FBgn0086356	lethal
Woma_00010706	FBgn0033699	lethal
Woma_00011083	FBgn0036695	partially lethal
Woma_00013747	FBgn0011787	lethal
Woma_00012282	FBgn0039680	partially lethal
Woma_00010967	FBgn0051852	lethal
Woma_00012366	FBgn0043012	lethal
Woma_00003936	FBgn0011715	lethal
Woma_00009627	FBgn0038106	lethal
Woma_00011969	FBgn0010078	lethal
Woma_00011957	FBgn0034012	lethal
Woma_00007584	FBgn0035025	lethal
Woma_00011511	FBgn0263831	partially lethal
Woma_00011089	FBgn0052302	partially lethal
Woma_00011190	FBgn0025615	partially lethal
Woma_00013479	FBgn0083946	lethal
Woma_00010143	FBgn0037892	lethal
Woma_00010161	FBgn0039627	lethal
Woma_00003106	FBgn0001970	lethal
Woma_00012680	FBgn0010292	lethal
Woma_00006889	FBgn0022774	lethal
Woma_00011376	FBgn0259676	lethal
Woma_00010950	FBgn0032209	lethal
Woma_00006110	FBgn0086451	lethal
Woma_00005426	FBgn0033763	lethal
Woma_00011513	FBgn0035524	lethal
Woma_00010846	FBgn0030091	lethal
Woma_00001288	FBgn0030645	partially lethal
Woma_00007277	FBgn0261789	lethal
Woma_00003807	FBgn0022943	lethal
Woma_00009337	FBgn0034354	lethal
Woma_00013278	FBgn0010620	lethal
Woma_00003983	FBgn0038706	lethal
Woma_00015260	FBgn0015615	lethal
Woma_00001187	FBgn0034029	partially lethal
Woma_00010620	FBgn0028375	lethal
Woma_00010311	FBgn0036133	partially lethal
Woma_00007865	FBgn0036356	lethal

Woma_00013052	FBgn0033128	partially lethal
Woma_00003162	FBgn0000575	lethal
Woma_00001771	FBgn0037834	partially lethal
Woma_00012469	FBgn0263607	lethal
Woma_00001482	FBgn0267823	lethal
Woma_00004182	FBgn0040985	partially lethal
Woma_00005565	FBgn0038140	lethal
Woma_00001796	FBgn0262801	lethal
Woma_00009918	FBgn0027329	lethal
Woma_00001269	FBgn0029992	lethal
Woma_00003871	FBgn0039407	lethal
Woma_00004021	FBgn0015399	lethal
Woma_00006917	FBgn0015828	lethal
Woma_00007636	FBgn0034109	lethal
Woma_00007539	FBgn0011704	lethal
Woma_00001198	FBgn0011297	lethal
Woma_00007698	FBgn0264922	lethal
Woma_00009876	FBgn0021742	lethal
Woma_00001852	FBgn0039260	lethal
Woma_00003464	FBgn0037875	lethal
Woma_00013158	FBgn0013432	partially lethal
Woma_00003305	FBgn0035722	partially lethal
Woma_00006006	FBgn0015270	lethal
Woma_00004905	FBgn0027592	lethal
Woma_00002596	FBgn0250791	lethal
Woma_00007337	FBgn0263782	lethal
Woma_00011528	FBgn0036300	lethal
Woma_00013396	FBgn0032420	partially lethal
Woma_00008279	FBgn0033995	lethal
Woma_00015336	FBgn0004908	lethal
Woma_00002708	FBgn0046763	lethal
Woma_00012716	FBgn0027553	partially lethal
Woma_00003672	FBgn0036282	lethal
Woma_00000940	FBgn0026369	lethal
Woma_00004208	FBgn0004567	lethal
Woma_00011332	FBgn0033153	lethal
Woma_00010754	FBgn0287768	lethal
Woma_00009324	FBgn0015797	lethal
Woma_00008754	FBgn0000115	lethal
Woma_00010198	FBgn0041147	lethal
Woma_00004199	FBgn0028398	partially lethal
Woma_00012158	FBgn0000520	lethal
Woma_00001397	FBgn0034065	lethal
Woma_00008827	FBgn0030407	lethal
Woma_00010678	FBgn0032262	partially lethal
Woma_00012203	FBgn0027617	lethal
Woma_00000209	FBgn0003865	lethal
Woma_00005621	FBgn0000492	lethal
Woma_00009820	FBgn0039805	partially lethal
Woma_00010831	FBgn0023514	partially lethal
Woma_00003051	FBgn0032819	partially lethal
Woma_00004159	FBgn0284084	lethal
Woma_00010725	FBgn0032020	partially lethal
Woma_00003481	FBgn0262559	lethal
Woma_00002110	FBgn0067102	partially lethal
Woma_00012300	FBgn0025832	lethal
Woma_00009229	FBgn0040268	lethal
Woma_00000882	FBgn0037610	partially lethal

Woma_00008291	FBgn0001148	lethal
Woma_00003512	FBgn0014366	lethal
Woma_00007399	FBgn0038045	partially lethal
Woma_00001564	FBgn0030317	partially lethal
Woma_00001892	FBgn0261065	lethal
Woma_00009377	FBgn0015929	lethal
Woma_00001676	FBgn0028331	lethal
Woma_00006695	FBgn0029885	lethal
Woma_00011875	FBgn0030521	lethal
Woma_00006167	FBgn0034909	partially lethal
Woma_00012779	FBgn0286213	lethal
Woma_00001095	FBgn0000395	lethal
Woma_00001802	FBgn0038118	lethal
Woma_00000792	FBgn0046685	partially lethal
Woma_00001980	FBgn0004391	lethal
Woma_00003210	FBgn0025355	lethal
Woma_00009371	FBgn0033179	partially lethal
Woma_00013719	FBgn0250850	partially lethal
Woma_00003845	FBgn0040336	lethal
Woma_00009650	FBgn0034994	partially lethal
Woma_00002842	FBgn0038578	lethal
Woma_00013614	FBgn0051803	partially lethal
Woma_00007138	FBgn0036886	lethal
Woma_00004651	FBgn0035718	partially lethal
Woma_00009666	FBgn0030855	lethal
Woma_00006096	FBgn0003866	lethal
Woma_00001963	FBgn0024996	lethal
Woma_00010339	FBgn0037249	lethal
Woma_00007225	FBgn0038964	lethal
Woma_00010647	FBgn0027868	lethal
Woma_00001613	FBgn0001612	lethal
Woma_00007823	FBgn0250848	lethal
Woma_00001207	FBgn0015372	lethal
Woma_00002167	FBgn0261245	lethal
Woma_00005581	FBgn0266268	lethal
Woma_00003211	FBgn0051251	partially lethal
Woma_00010597	FBgn0001086	lethal
Woma_00002662	FBgn0003277	lethal
Woma_00013255	FBgn0025186	lethal
Woma_00007096	FBgn0010830	lethal
Woma_00010895	FBgn0020497	lethal
Woma_00002749	FBgn0037293	lethal
Woma_00014065	FBgn0023542	lethal
Woma_00003204	FBgn0015296	partially lethal
Woma_00008597	FBgn0036000	partially lethal
Woma_00008574	FBgn0037073	lethal
Woma_00002120	FBgn0028292	lethal
Woma_00012875	FBgn0039626	lethal
Woma_00007937	FBgn0283536	lethal
Woma_00006980	FBgn0028978	partially lethal
Woma_00001853	FBgn0015240	lethal
Woma_00009676	FBgn0010391	lethal
Woma_00016239	FBgn0287185	lethal
Woma_00002839	FBgn0037933	lethal
Woma_00014687	FBgn0024986	lethal
Woma_00005224	FBgn0002354	lethal
Woma_00002633	FBgn0025640	partially lethal
Woma_00008985	FBgn0028695	lethal

Woma_00006857	FBgn0029714	lethal
Woma_00002985	FBgn0011232	partially lethal
Woma_00007329	FBgn0037574	lethal
Woma_00001058	FBgn0033741	lethal
Woma_00014551	FBgn0039566	lethal
Woma_00005991	FBgn0038853	lethal
Woma_00006831	FBgn0029733	lethal
Woma_00000962	FBgn0042180	lethal
Woma_00001845	FBgn0011703	lethal
Woma_00001925	FBgn0027093	lethal
Woma_00013806	FBgn0051217	lethal
Woma_00005094	FBgn0024230	partially lethal
Woma_00003873	FBgn0039402	lethal
Woma_00014758	FBgn0035975	partially lethal
Woma_00015122	FBgn0263109	partially lethal
Woma_00003981	FBgn0086613	lethal
Woma_00012783	FBgn0036916	partially lethal
Woma_00011561	FBgn0039828	lethal
Woma_00013222	FBgn0037121	lethal
Woma_00007926	FBgn0038746	lethal
Woma_00013808	FBgn0001112	lethal
Woma_00008171	FBgn0285910	lethal
Woma_00000660	FBgn0051065	lethal
Woma_00006767	FBgn0030434	lethal
Woma_00002538	FBgn0030612	lethal
Woma_00004692	FBgn0031947	partially lethal
Woma_00009851	FBgn0004597	lethal
Woma_00003989	FBgn0027560	lethal
Woma_00005349	FBgn0037671	lethal
Woma_00011473	FBgn0004643	lethal
Woma_00010071	FBgn0036598	lethal
Woma_00007081	FBgn0002573	lethal
Woma_00000439	FBgn0040512	lethal
Woma_00002454	FBgn0000137	lethal
Woma_00007349	FBgn0278604	lethal
Woma_00005080	FBgn0024191	partially lethal
Woma_00005253	FBgn0260962	lethal
Woma_00010909	FBgn0032030	lethal
Woma_00010755	FBgn0011294	partially lethal
Woma_00001854	FBgn0039259	partially lethal
Woma_00003818	FBgn0038275	lethal
Woma_00007614	FBgn0033813	lethal
Woma_00009761	FBgn0025743	partially lethal
Woma_00014946	FBgn0036213	lethal
Woma_00005660	FBgn0000576	lethal
Woma_00000776	FBgn0039233	lethal
Woma_00011457	FBgn0038467	lethal
Woma_00003964	FBgn0038692	lethal
Woma_00002418	FBgn0262685	lethal
Woma_00010946	FBgn0002021	lethal
Woma_00008639	FBgn0000715	lethal
Woma_00007879	FBgn0015075	lethal
Woma_00007878	FBgn0037156	partially lethal
Woma_00007755	FBgn0030569	lethal
Woma_00005955	FBgn0038595	partially lethal
Woma_00016452	FBgn0039907	lethal
Woma_00010347	FBgn0037778	lethal
Woma_00011592	FBgn0023129	lethal

Woma_00002058	FBgn0030969	lethal
Woma_00010704	FBgn0031607	partially lethal
Woma_00008584	FBgn0011016	lethal
Woma_00005393	FBgn0265351	lethal
Woma_00012261	FBgn0033554	lethal
Woma_00006396	FBgn0260756	lethal
Woma_00013090	FBgn0262601	lethal
Woma_00003273	FBgn0286567	lethal
Woma_00001026	FBgn0003391	lethal
Woma_00002626	FBgn0030088	lethal
Woma_00007387	FBgn0003512	lethal
Woma_00005783	FBgn0031247	lethal
Woma_00005150	FBgn0031802	lethal
Woma_00000552	FBgn0261274	lethal
Woma_00002911	FBgn0031851	lethal
Woma_00009406	FBgn0034564	lethal
Woma_00004633	FBgn0284253	lethal
Woma_00000871	FBgn0038146	partially lethal
Woma_00007497	FBgn0243511	lethal
Woma_00000909	FBgn0039509	lethal
Woma_00005519	FBgn0037797	partially lethal
Woma_00011678	FBgn0023167	lethal
Woma_00002537	FBgn0030611	lethal
Woma_00014248	FBgn0030367	lethal
Woma_00000580	FBgn0002284	lethal
Woma_00014747	FBgn0029944	lethal
Woma_00005961	FBgn0037328	lethal
Woma_00007358	FBgn0029503	lethal
Woma_00013265	FBgn0033526	lethal
Woma_00006588	FBgn0040477	lethal
Woma_00011212	FBgn0004053	lethal
Woma_00013198	FBgn0026320	lethal
Woma_00012315	FBgn0025879	partially lethal
Woma_00002116	FBgn0027334	lethal
Woma_00003374	FBgn0289108	lethal
Woma_00013612	FBgn0032634	lethal
Woma_00012611	FBgn0032705	lethal
Woma_00009093	FBgn0035676	partially lethal
Woma_00001684	FBgn0000140	lethal
Woma_00006910	FBgn0261109	lethal
Woma_00006321	FBgn0266669	lethal
Woma_00007679	FBgn0020391	lethal
Woma_00007745	FBgn0033206	lethal
Woma_00008233	FBgn0030816	lethal
Woma_00007393	FBgn0029512	lethal
Woma_00013733	FBgn0033235	partially lethal
Woma_00004983	FBgn0027779	lethal
Woma_00004638	FBgn0027609	partially lethal
Woma_00008113	FBgn0030545	lethal
Woma_00001195	FBgn0010238	lethal
Woma_00007472	FBgn0038331	lethal
Woma_00014064	FBgn0003116	lethal
Woma_00005421	FBgn0021825	lethal
Woma_00007486	FBgn0039008	partially lethal
Woma_00001657	FBgn0029665	lethal
Woma_00002422	FBgn0001491	lethal
Woma_00001212	FBgn0264339	partially lethal
Woma_00014520	FBgn0261850	lethal

Woma_00006501	FBgn0034943	partially lethal
Woma_00005761	FBgn0011638	lethal
Woma_00007727	FBgn0261524	lethal
Woma_00011174	FBgn0029629	lethal
Woma_00002788	FBgn0010421	lethal
Woma_00005582	FBgn0039882	lethal
Woma_00016058	FBgn0037884	partially lethal
Woma_00011011	FBgn0027091	lethal
Woma_00009816	FBgn0003720	lethal
Woma_00000911	FBgn0004644	lethal
Woma_00011758	FBgn0033688	lethal
Woma_00003935	FBgn0037330	lethal
Woma_00009633	FBgn0034243	lethal
Woma_00000619	FBgn0030551	partially lethal
Woma_00006861	FBgn0026432	partially lethal
Woma_00002446	FBgn0261284	lethal
Woma_00004384	FBgn0015282	lethal
Woma_00014208	FBgn0036822	partially lethal
Woma_00009617	FBgn0002736	lethal
Woma_00011865	FBgn0286027	lethal
Woma_00002069	FBgn0035458	partially lethal
Woma_00010197	FBgn0035294	lethal
Woma_00009318	FBgn0286781	lethal
Woma_00010685	FBgn0032919	lethal
Woma_00000964	FBgn0033482	partially lethal
Woma_00011663	FBgn0023181	partially lethal
Woma_00005931	FBgn0038868	lethal
Woma_00012608	FBgn0260861	lethal
Woma_00009028	FBgn0035401	lethal
Woma_00012623	FBgn0015561	lethal
Woma_00015333	FBgn0264326	lethal
Woma_00007931	FBgn0038752	partially lethal
Woma_00011189	FBgn0029718	partially lethal
Woma_00012982	FBgn0025571	lethal
Woma_00006518	FBgn0020764	lethal
Woma_00009632	FBgn0040294	lethal
Woma_00001158	FBgn0260456	lethal
Woma_00001733	FBgn0039067	lethal
Woma_00014253	FBgn0037270	lethal
Woma_00010880	FBgn0015271	lethal
Woma_00011295	FBgn0270928	lethal
Woma_00013451	FBgn0039195	lethal
Woma_00006411	FBgn0000543	lethal
Woma_00010808	FBgn0026679	lethal
Woma_00013164	FBgn0014411	partially lethal
Woma_00009885	FBgn0003366	lethal
Woma_00013395	FBgn0032421	lethal
Woma_00005287	FBgn0039155	partially lethal
Woma_00013539	FBgn0038968	lethal
Woma_00012695	FBgn0030067	lethal
Woma_00006965	FBgn0259791	lethal
Woma_00013600	FBgn0000061	lethal
Woma_00005833	FBgn0032394	lethal
Woma_00009605	FBgn0263510	lethal
Woma_00003681	FBgn0035435	lethal
Woma_00004014	FBgn0031538	partially lethal
Woma_00012070	FBgn0033400	lethal
Woma_00014251	FBgn0028426	lethal

Woma_00009083	FBgn0035627	lethal
Woma_00001933	FBgn0261064	lethal
Woma_00013813	FBgn0037489	lethal
Woma_00012469	FBgn0263606	lethal
Woma_00000463	FBgn0020616	lethal
Woma_00006471	FBgn0034232	lethal
Woma_00007240	FBgn0027493	lethal
Woma_00008032	FBgn0035192	partially lethal
Woma_00014752	FBgn0023458	lethal
Woma_00000705	FBgn0017590	lethal
Woma_00002030	FBgn0026323	lethal
Woma_00008269	FBgn0003892	lethal
Woma_00004450	FBgn0028738	lethal
Woma_00003006	FBgn0011761	lethal
Woma_00011914	FBgn0032929	lethal
Woma_00003010	FBgn0003046	partially lethal
Woma_00001785	FBgn0261108	lethal
Woma_00002220	FBgn0037478	lethal
Woma_00000826	FBgn0020018	lethal
Woma_00006786	FBgn0034537	lethal
Woma_00011285	FBgn0030063	lethal
Woma_00007189	FBgn0015295	lethal
Woma_00008825	FBgn0000316	lethal
Woma_00012311	FBgn0039702	partially lethal
Woma_00009362	FBgn0263260	lethal
Woma_00010355	FBgn0039469	partially lethal
Woma_00009056	FBgn0025712	lethal
Woma_00006886	FBgn0036514	lethal
Woma_00010155	FBgn0036973	lethal
Woma_00011904	FBgn0015391	lethal
Woma_00002114	FBgn0021874	lethal
Woma_00003904	FBgn0002633	lethal
Woma_00012442	FBgn0086357	lethal
Woma_00012717	FBgn0022786	lethal
Woma_00004094	FBgn0031959	lethal
Woma_00005750	FBgn0003687	lethal
Woma_00006315	FBgn0015550	lethal
Woma_00014684	FBgn0023535	partially lethal
Woma_00013298	FBgn0264294	lethal
Woma_00014236	FBgn0086785	lethal
Woma_00002804	FBgn0038454	lethal
Woma_00015350	FBgn0037231	partially lethal
Woma_00003041	FBgn0032514	lethal
Woma_00010829	FBgn0040394	lethal
Woma_00005448	FBgn0033961	lethal
Woma_00007238	FBgn0019936	lethal
Woma_00012617	FBgn0032690	lethal
Woma_00007847	FBgn0036990	lethal
Woma_00011508	FBgn0024285	lethal
Woma_00010128	FBgn0051155	lethal
Woma_00000346	FBgn0035736	partially lethal
Woma_00007177	FBgn0000709	lethal
Woma_00013063	FBgn0027558	lethal
Woma_00013179	FBgn0023527	lethal
Woma_00011753	FBgn0034853	lethal
Woma_00001747	FBgn0037240	lethal
Woma_00015964	FBgn0031020	lethal
Woma_00003447	FBgn0037856	lethal

Woma_00000051	FBgn0029818	lethal
Woma_00005886	FBgn0019932	lethal
Woma_00004753	FBgn0031627	lethal
Woma_00006511	FBgn0040273	lethal
Woma_00005271	FBgn0264325	lethal
Woma_00003188	FBgn0261799	lethal
Woma_00001205	FBgn0028983	lethal
Woma_00009403	FBgn0008636	lethal
Woma_00015338	FBgn0037741	lethal
Woma_00008451	FBgn0003255	lethal
Woma_00002997	FBgn0001941	lethal
Woma_00006404	FBgn0052251	lethal
Woma_00005444	FBgn0026378	lethal
Woma_00004131	FBgn0031304	partially lethal
Woma_00003833	FBgn0014879	partially lethal
Woma_00003744	FBgn0037882	lethal
Woma_00013759	FBgn0036742	partially lethal
Woma_00002879	FBgn0037301	lethal
Woma_00004510	FBgn0262560	lethal
Woma_00012844	FBgn0037220	partially lethal
Woma_00006393	FBgn0035838	lethal
Woma_00007079	FBgn0286818	lethal
Woma_00002036	FBgn0015789	partially lethal
Woma_00000745	FBgn0037783	lethal
Woma_00007011	FBgn0052227	partially lethal
Woma_00002051	FBgn0023506	partially lethal
Woma_00002931	FBgn0001977	lethal
Woma_00009713	FBgn0052645	lethal
Woma_00005136	FBgn0263278	lethal
Woma_00000809	FBgn0039728	partially lethal
Woma_00003184	FBgn0027080	lethal
Woma_00000443	FBgn0028992	lethal
Woma_00001011	FBgn0003527	lethal
Woma_00011453	FBgn0266111	lethal
Woma_00003665	FBgn0036277	lethal
Woma_00010201	FBgn0025682	lethal
Woma_00005556	FBgn0038130	lethal
Woma_00010908	FBgn0261437	lethal
Woma_00013354	FBgn0085484	lethal
Woma_00005835	FBgn0032859	partially lethal
Woma_00009801	FBgn0023179	lethal
Woma_00004079	FBgn0024291	lethal
Woma_00009054	FBgn0010894	lethal
Woma_00009382	FBgn0286788	lethal
Woma_00012722	FBgn0030141	lethal
Woma_00009399	FBgn0034558	partially lethal
Woma_00006491	FBgn0011726	lethal
Woma_00000076	FBgn0008635	lethal
Woma_00015274	FBgn0011573	lethal
Woma_00009782	FBgn0010380	lethal
Woma_00003853	FBgn0038772	lethal
Woma_00010969	FBgn0040228	lethal
Woma_00008208	FBgn0030081	lethal
Woma_00004858	FBgn0031740	lethal
Woma_00006170	FBgn0034913	lethal
Woma_00012365	FBgn0285937	lethal
Woma_00010150	FBgn0266195	lethal
Woma_00000380	FBgn0010747	lethal

Woma_00008235	FBgn0052708	lethal
Woma_00012281	FBgn0033169	partially lethal
Woma_00005428	FBgn0033454	lethal
Woma_00005404	FBgn0003114	lethal
Woma_00002048	FBgn0014464	partially lethal
Woma_00007140	FBgn0262517	lethal
Woma_00011367	FBgn0014269	lethal
Woma_00010745	FBgn0022213	lethal
Woma_00013488	FBgn0039283	lethal
Woma_00002714	FBgn0037351	lethal
Woma_00011991	FBgn0033911	lethal
Woma_00000065	FBgn0030529	lethal
Woma_00000739	FBgn0039489	lethal
Woma_00004166	FBgn0263133	lethal
Woma_00004822	FBgn0010660	lethal
Woma_00010753	FBgn0001983	lethal
Woma_00010833	FBgn0001259	lethal
Woma_00006395	FBgn0029113	lethal
Woma_00012775	FBgn0036487	lethal
Woma_00010331	FBgn0038923	partially lethal
Woma_00003533	FBgn0035699	lethal
Woma_00011039	FBgn0033349	lethal
Woma_00007948	FBgn0266672	lethal
Woma_00009722	FBgn0022724	lethal
Woma_00006239	FBgn0034915	lethal
Woma_00003189	FBgn0004569	lethal
Woma_00013887	FBgn0027783	lethal
Woma_00012810	FBgn0036448	lethal
Woma_00004212	FBgn0027363	lethal
Woma_00010034	FBgn0283681	lethal
Woma_00004931	FBgn0010520	lethal
Woma_00011537	FBgn0015919	lethal
Woma_00005226	FBgn0038055	partially lethal
Woma_00013463	FBgn0039404	lethal
Woma_00011503	FBgn0003330	lethal
Woma_00012983	FBgn0010926	lethal
Woma_00000116	FBgn0052793	lethal
Woma_00009947	FBgn0030410	lethal
Woma_00005694	FBgn0004863	lethal
Woma_00001784	FBgn0038989	lethal
Woma_00009065	FBgn0036124	lethal
Woma_00009970	FBgn0052344	lethal
Woma_00003463	FBgn0037874	lethal
Woma_00004641	FBgn0031606	partially lethal
Woma_00003067	FBgn0011291	partially lethal
Woma_00015829	FBgn0039914	partially lethal
Woma_00007128	FBgn0036314	lethal
Woma_00011527	FBgn0036299	lethal
Woma_00011436	FBgn0027945	lethal
Woma_00009353	FBgn0034602	lethal
Woma_00011304	FBgn0039461	partially lethal
Woma_00007330	FBgn0062412	lethal
Woma_00012654	FBgn0033728	lethal
Woma_00006606	FBgn0033855	lethal
Woma_00003108	FBgn0261535	lethal
Woma_00005987	FBgn0028468	lethal
Woma_00011202	FBgn0001077	lethal
Woma_00001734	FBgn0038617	partially lethal

Woma_00001418	FBgn0033210	lethal
Woma_00000833	FBgn0039218	lethal
Woma_00004509	FBgn0001308	lethal
Woma_00007816	FBgn0036804	lethal
Woma_00008905	FBgn0030717	partially lethal
Woma_00003808	FBgn0039204	partially lethal
Woma_00015672	FBgn0266671	lethal
Woma_00013130	FBgn0260817	lethal
Woma_00001578	FBgn0000499	lethal
Woma_00005289	FBgn0031114	lethal
Woma_00004913	FBgn0031254	partially lethal
Woma_00009652	FBgn0050419	lethal
Woma_00015849	FBgn0041723	lethal
Woma_00007321	FBgn0039227	lethal
Woma_00006378	FBgn0003515	lethal
Woma_00000801	FBgn0037815	partially lethal
Woma_00008768	FBgn0034826	partially lethal
Woma_00005097	FBgn0259683	partially lethal
Woma_00005651	FBgn0261434	lethal
Woma_00004217	FBgn0000588	lethal
Woma_00003256	FBgn0011591	partially lethal
Woma_00003743	FBgn0261575	lethal
Woma_00005284	FBgn0266673	lethal
Woma_00004832	FBgn0000221	lethal
Woma_00011555	FBgn0285911	lethal
Woma_00000474	FBgn0035370	lethal
Woma_00002614	FBgn0030352	lethal
Woma_00002870	FBgn0261984	lethal
Woma_00011856	FBgn0022224	lethal
Woma_00010021	FBgn0264848	lethal
Woma_00009409	FBgn0050342	lethal
Woma_00000561	FBgn0035511	partially lethal
Woma_00000845	FBgn0260779	lethal
Woma_00002062	FBgn0026411	lethal
Woma_00007135	FBgn0024889	lethal
Woma_00003377	FBgn0259113	lethal
Woma_00001783	FBgn0014949	lethal
Woma_00011864	FBgn0022069	lethal
Woma_00008545	FBgn0045759	lethal
Woma_00008262	FBgn0034734	lethal
Woma_00013127	FBgn0010356	lethal
Woma_00006165	FBgn0028411	partially lethal
Woma_00010470	FBgn0034084	lethal
Woma_00001398	FBgn0010590	lethal
Woma_00010910	FBgn0002022	lethal
Woma_00005227	FBgn0038320	partially lethal
Woma_00010691	FBgn0032925	lethal
Woma_00004602	FBgn0033379	lethal
Woma_00010257	FBgn0002926	lethal
Woma_00008209	FBgn0032346	partially lethal
Woma_00010649	FBgn0032292	lethal
Woma_00007392	FBgn0086134	lethal
Woma_00009604	FBgn0033450	partially lethal
Woma_00005912	FBgn0039212	lethal
Woma_00008709	FBgn0003977	lethal
Woma_00003931	FBgn0000147	lethal
Woma_00004092	FBgn0015905	lethal
Woma_00001013	FBgn0016970	lethal

Woma_00009051	FBgn0036733	lethal
Woma_00001263	FBgn0032244	lethal
Woma_00007627	FBgn0025716	lethal
Woma_00010404	FBgn0028687	lethal
Woma_00007454	FBgn0037608	lethal
Woma_00005219	FBgn0038678	lethal
Woma_00010220	FBgn0035471	lethal
Woma_00006048	FBgn0038309	lethal
Woma_00006486	FBgn0040465	partially lethal
Woma_00008329	FBgn0036500	lethal
Woma_00008755	FBgn0263600	lethal
Woma_00010137	FBgn0027095	lethal
Woma_00003070	FBgn0032144	lethal
Woma_00003031	FBgn0029755	lethal
Woma_00004690	FBgn0005390	lethal
Woma_00004123	FBgn0032703	partially lethal
Woma_00003491	FBgn0037660	partially lethal
Woma_00013276	FBgn0021795	lethal
Woma_00001034	FBgn0022700	partially lethal
Woma_00002427	FBgn0052499	lethal
Woma_00012173	FBgn0039306	lethal
Woma_00005597	FBgn0024841	lethal
Woma_00002451	FBgn0029925	partially lethal
Woma_00007125	FBgn0030359	lethal
Woma_00010856	FBgn0035194	lethal
Woma_00001191	FBgn0028683	lethal
Woma_00002375	FBgn0029525	lethal
Woma_00003801	FBgn0037708	partially lethal
Woma_00004117	FBgn0015834	lethal
Woma_00008631	FBgn0053349	lethal
Woma_00005389	FBgn0010433	lethal
Woma_00005681	FBgn0040623	lethal
Woma_00006126	FBgn0032154	lethal
Woma_00004066	FBgn0032728	lethal
Woma_00012352	FBgn0039527	lethal
Woma_00007299	FBgn0027527	lethal
Woma_00008330	FBgn0038815	lethal
Woma_00003480	FBgn0011701	lethal
Woma_00008548	FBgn0001104	lethal
Woma_00002732	FBgn0003002	lethal
Woma_00001469	FBgn0034073	partially lethal
Woma_00009778	FBgn0031062	lethal
Woma_00000782	FBgn0010438	lethal
Woma_00000912	FBgn0086359	lethal
Woma_00012543	FBgn0010638	lethal
Woma_00014585	FBgn0262512	lethal
Woma_00002690	FBgn0014861	lethal

7. DISCUSSION & CONCLUSION

Myiasis still remains an unresolved problem for animal production in the countries of the Global South as well as in Western countries (Francesconi and Lupi, 2012; Hall et al., 2016; Zumpt, 1965). Specifically, traumatic myiasis caused by three fly species—*L. cuprina*, *C. hominivorax*, and *W. magnifica*—severely harms the animal husbandry industry (Hall et al., 2016). *L. cuprina* is prevalent in Australia and New Zealand, earning it the notorious moniker "Australian sheep blowfly" due to its profound impact on the Australian and New Zealand sheep industry (Kotze and James, 2022; Sandeman et al., 2014). *C. hominivorax* is endemic in South America and has been successfully eradicated from North and Central America through SIT (Knipling, 1955; Scott et al., 2017; Wyss, 2000). In contrast, *W. magnifica* has a significant impact on mammals across Asia, Africa, and Europe. However, both fundamental and applied studies lag far behind in this species. In this thesis, we conducted genomic and transcriptomic studies to fill these gaps.

7.1 The genome and evolutionary relationship of W. magnifica

Chromosome-level assembled and well-annotated reference genome resources are the starting point for many fundamental and applied studies on myiasis-causing flies. Such investigations include the identification of the genes that enable myiasis-causing flies to parasitize their hosts, the understanding of the mechanisms of insecticide resistance at the molecular level, the search for novel vaccine and insecticide targets, and the isolation of key genes involved in the development of genetic control measures. The genome of *L. cuprina* was sequenced by Anstead et al. (2015) and, that of *C. hominivorax* by Scott et al. (2020). In order to reduce variability, the nucleic acid samples used for sequencing were extracted from highly inbred strains in both species. However, the stable laboratory rearing of *W. magnifica* remains elusive, posing a challenge for acquiring the inbred flies. If non-inbred samples are utilized for genome sequencing, it may present several drawbacks, such as increasing the complexities of genomic data, resulting in greater overlap and fragmentation during the assembly phase, demanding increased efforts in discerning gene locations and functions during annotation, and necessitating a greater allocation of computational resources and a longer time frame. In the first article, we adopted the strategy of low DNA input library preparation (Kingan, et al., 2019) and successfully sequenced the genome of *W. magnifica* from a single adult fly. The results

from Benchmarking Universal Single-Copy Orthologs (BUSCO) (Simão et al., 2015) and other assessment criteria indicated the high quality of the genome, with only a limited number of missing genes, on par with genomes sequenced using other approaches (Anstead et al., 2015; Scott et al., 2020). The strategy employed in this study can be established as a benchmark sequencing approach for certain small dipterous flies that are difficult to rear in the laboratory, or are suitable for lab rearing but prove difficult to create inbred lines.

The genome of *W. magnifica* was 753.99 Mb (N50: 5.00 Mb; GC content: 32.82%). In comparison with other Diptera flies, this is larger than *Sarcophaga bullata* (Martinson et al., 2019), *L. cuprina* (Anstead et al., 2015), and *C. hominivorax* (Scott et al., 2020), and even is five times larger than *Drosophila melanogaster* (Adams et al., 2000). The genome of *W. magnifica* contains significantly higher amounts of repetitive elements than those in the four Diptera genomes mentioned above, assuming that a larger portion of the genome may be mainly derived from repetitive sequences. We predicted 16,718 protein-coding genes in the genome of *W. magnifica*. These annotated genes are of great importance for the investigation of the gene expression profiles at different developmental stages (as performed in the second article) and for the identification of the elements for the development of genetic control methods (as described in the third article). Interestingly, 2062 (10.37%) genes were predicted to be unique to *W. magnifica* (no homologous genes were found in any other selected Diptera), which may be involved in the *W. magnifica*-specific parasitic activities.

The evolutionary analysis revealed that *W. magnifica* and *S. bullata* are closely related, followed by *L. cuprina*, while *A. aegypti* and *A. gambiae* are the most distantly related. This is in line with previous taxonomic classifications, as *S. bullata* also belongs to the family Sarcophagidae, while *L. cuprina* is another myiasis-causing fly in the family Calliphoridae, and *A. aegypti* and *A. gambiae* are categorized under the mosquito genera. In the third article, we also used the *W. magnifica's* transformer (*Wmtra*) and transformer2 (*Wmtra2*) genes to perform an evolutionary analysis with other Diptera, respectively and both results were also consistent with taxonomic classifications.

7.2 The investigation of the gene expression profiles

In the first article, gene family analysis revealed that the expanded gene families in *W. magnifica* are responsible for immunity, insecticide resistance mechanisms, heat stress

response, cuticle development, and transcription/chromatin modification. Additionally, an analysis of the genes specific to three myiasis-causing flies uncovered that the genes such as olfactory-related genes, insecticide resistance-related genes, and proteases, are likely related to parasitism. In the second article, 2049 excretory/secretory (ES) proteins closely associated with parasitism were identified in *W. magnifica*; their roles are linked to cuticle development, peptidase activity, immune responses, metabolic processes, and nutrient storage. Taking together, these genes are likely to have a central role in the successful parasitism of *W. magnifica* from first-stage larvae entering the host wounds to third-stage larvae leaving the host wounds to pupate.

In the second article, to gain a better understanding of how the larvae parasitize their hosts, we targeted the peptidases, cuticle proteins, heat shock proteins, and immune response genes and investigated the expression of these parasitism-related genes in larvae collected from host wounds in comparison to other following stages. In the beginning, when adult female flies lay first-stage larvae in infestation sites on hosts, the larvae need to penetrate the host's skin to invade tissues and feed on it. In *L. cuprina* and *O. ovis*, the invasion and nutrient digestion from the host's skin are largely dependent on proteases (Bowles et al., 1988; Casu et al., 1994, 1996; Sandeman et al., 1990; Tabouret et al., 2003). During the parasitic larval stage of *W. magnifica*, a significant number of peptidases were highly expressed compared to other developmental stages. In particular, for serine peptidases that are closely associated with larval parasitism, up to 22.93% (47/205) were highly expressed during the parasitic larval stages, in comparison, only 4.39% (9/205) during the pupal stage, and 8.78% (18/205) during the adult stage. This suggests that during the larval stage, the upregulation of peptidases, especially serine peptidases, is crucial for tissue invasion and nutrient acquisition from the hosts. Cuticles have a variety of key functions in the biology of insects, not only structuring their tough exoskeletons, but also serving as a barrier between living tissues and the external environment to protect them from dehydration, mechanical injury and predation, and insecticides (Andersen, 1979; Balabanidou et al., 2018; Muthukrishnan et al., 2020). Notably, over half (51.16%) of the cuticle protein genes are highly expressed during the parasitic larval stages of *W. magnifica*. This indicates that the cuticle proteins, on the one hand, are involved in the molting process, adapting to the increase in size of larvae of *W. magnifica* and on the other hand, likely protect the parasitic larvae from harm from the hosts or the external environment. Traumatic myiasis can harm host tissues and make wounds prone to bacterial infection (Erdmann, 1987). As larvae enter the host wounds, they must activate their immune response to defend against bacteria,

which includes the release of antimicrobial peptides (AMPs) (Yi et al., 2014). In *W. magnifica*, 1 defensin and 1 diptericin AMP genes exhibited higher expression in second-stage larvae, which likely protects larvae from both Gram-positive and Gram-negative bacteria infection in the host wound. Heat shock proteins (hsp) are produced in cells under high-temperature or stress conditions, safeguarding cells from damage (King and MacRae, 2015). We noticed the increased expression of 21 hsp genes during the larval stages. This is likely because the body temperature of camels is higher than the optimal development temperature of *W. magnifica*. Furthermore, the genes linked to nutrient reservoir activity are predominantly highly expressed in third-stage larvae. The nutrient reserves are vital for the pupal and adult stages. For example, in *O. ovis*, a low larval body weight (below 280 mg) can impact survival in the later stages (Cepeda-Palacios et al., 2000). The strong expression of the genes indicates that nutrient accumulation occurs during the third larval stage.

In the second article, we also investigated all gene expression profiles to reveal the details of the gene regulatory mechanisms during the life cycle of *W. magnifica* and to provide valuable information for the management of *W. magnifica*. Throughout the developmental stages of *W. magnifica*, a multitude of genes exhibit distinct activities: in the second larval stage, the genes related to cuticle proteins, peptidase activity, and RNA transcription and translation are prominent; moving into the third larval stage, the genes associated with peptidase inhibitor activity and nutrient reservoir activity take the lead; in the pupal stage, numerous genes that govern cell and tissue morphogenesis and cell and tissue development were upregulated; and in the adult stage, the upregulated genes play crucial roles in perception, particularly light perception, and adult behaviors including feeding, mating, and locomotion. The high expression of these specific genes at each developmental stage can be explained: in the second larval stage, with relatively rapid growth and development, the larvae secrete a large number of peptidases to break down host tissues for nutrient acquisition; in the third larval stage, the larvae start to accumulate nutrients for subsequent utilization of the pupae and adult flies; during the pupal stage, they undergo a significant morphological change and complete the transformation from the larval into the adult body; in the adult stage, adult flies adapt to the environment, regulate circadian rhythms, and engage in adult behaviors such as foraging for food.

7.3 The identification of key elements for the development of control approaches

Until now, insecticides have remained the most commonly used and effective method for treating myiasis. However, excessive use of insecticides can lead to the emergence of less susceptible individuals, resulting in the development of resistance. This resistance can diminish the long-term efficacy of insecticides, necessitating higher doses and further increasing both costs and environmental risks. For example, the myiasis-causing blow flies, *L. cuprina* and *C. hominivorax*, have developed resistance to various insecticides through the metabolic and target site insensitivity-resistance mechanisms (Carvalho et al., 2009; da Silva and de Azeredo-Espin, 2009; Sandeman et al., 2014; Tandonnet et al., 2022). While investigations into insecticide resistance in *W. magnifica* are currently lacking, the extensive historical use of insecticides for managing *W. magnifica* causing myiasis suggests the possible development of insecticide resistance. It is important to conduct systematic surveys to assess the resistance status of *W. magnifica*. As for vaccines, a great deal of work to develop a vaccine for *L. cuprina* was undertaken from the late 1980s until the early 2000s in Australia (Kotze and James, 2022; Sandeman et al., 2014), however, the research program ended with no vaccine being commercialized, while for *W. magnifica* no vaccines developed are reported. In the third article, we identified 972 genes that could potentially serve as potential targets for the development of insecticides or vaccines against *W. magnifica*.

Conditional female-lethal transgenic strains were developed in blow flies (Concha et al., 2016, 2020; Li et al., 2014; Yan and Scott, 2020), with the aim of avoiding the challenge of selecting the dose of radiation sterilization and saving the costs generated by females in traditional SIT. The key to making this system sex-specific is the *tra* gene. In *L. cuprina* and *C. hominivorax* the system was developed, the *tra* transcripts undergo sex-specific splicing, leading to a full-length, functional protein in females, while the male transcripts produce a truncated, non-functional polypeptide due to the male-specific exon inclusion, which contains multiple in-frame stop codons (Concha et al., 2009; Li et al., 2013). In the third article, we have successfully identified and characterized *Wmtra*. Similar to *L. cuprina* and *C. hominivorax*, the *Wmtra* transcripts exhibit a sex-specific splicing pattern. In our future work aimed at developing a conditional female-lethal transgene system in *W. magnifica*, the introduction of the crucial first intron of the *Wmtra* gene into the pro-apoptotic gene gives us an optimistic expectation that this system may perform effectively in *W. magnifica*, much as it has in other blow fly species like *L. cuprina* (Li et al., 2014; Yan and Scott, 2020) and *C. hominivorax* (Concha et al., 2016, 2020).

In terms of the genetic control of myiasis-causing flies, CRISPR/Cas9-based homing gene drive systems are expected to provide another powerful and promising tool. The system requires a cassette expressing a Cas9 and a gRNA that integrates into a genome at the precise site guided by the gRNA (Bier, 2022). Candidate sites for the introduction of the cassette into the genome are genes essential for female survival, development or fecundity. In addition, for Cas9 and gRNA to be expressed at an early stage, the identification of promoters from active genes in the germline (to drive Cas9 expression) and promoters from U6 RNA (to drive gRNA expression) is also essential. In the third article, we identified the female-biased and male-biased gene sets that can provide a range of useful target genes for the introduction of Cas9-gRNA cassette. Further, the female-biased gene set contains several genes whose promoters are frequently applied to drive Cas9 expression. Among them, based on the analysis of the *Wmnanos* gene and its successful application in *D. melanogaster* (Champer et al., 2022) and *D. suzukii* (Yadav et al., 2023), the promoter from *Wmnanos* is a better candidate to drive Cas9 expression for future work on a Cas9-based homing gene drive system for *W. magnifica*. In Cas9-based homing gene drive systems, promoters originating from the U6 RNA genes are well-documented for governing gRNA transcription in several mosquito species (Carballar-Lejarazú et al., 2020; Gantz et al., 2015; Hammond et al., 2016; Kyrou et al., 2018) and fly species (Champer et al., 2022; Yadav et al., 2023). We have successfully identified 6 U6 genes in the *W. magnifica* genome, which is a valuable resource for employing their promoters to facilitate gRNA expression.

7.4. Limitations and outlook of the thesis

In this thesis, we have employed a range of techniques to analyze *W. magnifica*, as described above. However, the fact that *W. magnifica* cannot be reared successfully in the laboratory has significantly limited the depth and width of the analysis. Therefore, my next priority is to explore adjustments to various conditions and diets in order to establish a stable laboratory-rearing condition for *W. magnifica*. In the first article, although I obtained a high-quality genome assembly with a high N50 value, it was at the contig level. Genome chromosome-level assemblies have significant advantages in deciphering genome structure and function, so in the next step, I will use Hi-C or Chicago technologies to obtain a chromosome-level genome assembly of *W. magnifica* and to identify the putative genes and genomic regions belonging to the sex chromosomes. In the second article, I successfully investigated larvae, pupae, and adult flies of *W. magnifica*, but other early developmental stages were not included. Next, with the

successful rearing in the laboratory of *W. magnifica*, I will investigate the gene expression patterns of other early developmental stages. Additionally, I will conduct a comparative analysis of differentially expressed genes between the parasitic larvae in the wound environment and the laboratory-reared larvae to gain a clearer understanding of the specific genes involved in the parasitism of *W. magnifica*. In the third article, I identified and characterized the key elements in the conditional female-lethal transgenic strain and the CRISPR/Cas9-based homing gene drive systems. However, their functions have not been experimentally validated. In the next step, I plan to employ techniques such as RNA interference (RNAi), CRISPR-Cas, or other methods to elucidate the functional roles of *tra*, *tra2*, the *nanos* gene promoter, and the U6 gene promoter in *W magnifica*. This will establish a robust foundation for the development of these two systems in *W. magnifica*.

Concluding remarks

In the first article, we introduced a strategy for low DNA input library preparation and successfully sequenced the genome of *W. magnifica* using a single fly specimen. This novel approach shows promise for its application to small Diptera flies that are challenging to rear in the laboratory or have difficulties with inbreeding. The genome of *W. magnifica* serves as a valuable resource for fundamental research in molecular biology, biochemistry, and genetics. Additionally, it provides opportunities for various applications in comparative genomics, transcriptomics, functional genomics, population genetics, and in particular, the development of novel control strategies. In the second article, we carried out a comprehensive exploration of the complex gene expression patterns throughout the life cycle of *W. magnifica*. The findings not only enhance our understanding of the molecular-level parasitic life cycle of *W. magnifica* but also lay the foundation for the development of insecticides or vaccines to disrupt larval parasitism on or in hosts. In the third article, we identified and characterized the essential components for the development of genetically modified strains in *W. magnifica*. This represents a crucial step towards establishing genetic control programs for managing the infestation of *W. magnifica*. In summary, this thesis has shed light on the biology of *W. magnifica* and has contributed to the development of innovative methods for controlling this myiasis-causing flesh fly, *W. magnifica*.

REFERENCES

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. Science. 2000;287(5461):2185-95.

Andersen SO. Biochemistry of insect cuticle. Annual Review of Entomology. 1979;24(1):29-59.

Ansari MS, Moraiet MA and Ahmad S. Insecticides: impact on the environment and human health. In: Malik A, Grohmann E, Akhtar R, editors. Environmental Deterioration and Human Health: Natural and Anthropogenic Determinants. Dordrecht: Springer; 2014. p. 99-123.

Anstead CA, Korhonen PK, Young ND, Hall RS, Jex AR, Murali SC, et al. *Lucilia cuprina* genome unlocks parasitic fly biology to underpin future interventions. Nature communications. 2015;6(1):7344.

An X, Yang B, Bao H, Oyun G, Wang X, Demtu E. Morphological observation of the larva of the alxa bactrian camel vaginal myiasis. Journal of Camel Practice and Research. 2019;26(1):57-62.

Balabanidou V, Grigoraki L, Vontas J. Insect cuticle: a critical determinant of insecticide resistance. Current Opinion in Insect Science. 2018;27:68-74.

Beyhan YE, YILMAZ H, CENGİZ ZT, Ayral A. A case of aural myiasis caused by *Wohlfahrtia magnifica* in a child in Turkey. Türk Hijyen ve Deneysel Biyoloji Dergisi. 2017;74(4):347-50.

Bier E. Gene drives gaining speed. Nature Reviews Genetics. 2022;23(1):5-22.

Boulard C. Degradation of bovine C3 by serine proteases from parasites *Hypoderma lineatum* (Diptera, Oestridae). Veterinary Immunology and Immunopathology. 1989;20(4):387-98.

Bowles VM, Carnegie PR, Sandeman RM. Characterization of proteolytic and collagenolytic enzymes from the larvae of *Lucilia cuprina*, the sheep blowfly. Australian Journal of Biological

Sciences. 1988;41(2):269-78.

Bushland, RC, Lindquist AW, Knipling EF. Eradication of screw-worms through release of sterilized males. Science. 1955;122:287-8.

Carballar-Lejarazú R, Ogaugwu C, Tushar T, Kelsey A, Pham TB, Murphy J, et al. Next-generation gene drive for population modification of the malaria vector mosquito, *Anopheles gambiae*. Proceedings of the National Academy of Sciences. 2020;117(37):22805-14.

Carnevali F, Franchini D, Otranto D, Giangaspero A, Di Bello A, Ciccarelli S, et al. A formulation of neem and hypericum oily extract for the treatment of the wound myiasis by *Wohlfahrtia magnifica* in domestic animals. Parasitology research. 2019;118:2361-7.

Carvalho RA, Torres TT, Paniago MG, Azeredo-Espin AML. Molecular characterization of esterase E3 gene associated with organophosphorus insecticide resistance in the New World screwworm fly, *Cochliomyia hominivorax*. Medical and Veterinary Entomology. 2009;23:86-91.

Casu RE, Eisemann CH, Vuocolo T, Tellam RL. The major excretory/secretory protease from *Lucilia cuprina* larvae is also a gut digestive protease. International Journal for Parasitology. 1996;26(6):623-8.

Casu RE, Jarmey JM, Elvin CM, Eisemann CH. Isolation of a trypsin-like serine protease gene family from the sheep blowfly *Lucilia cuprina*. Insect Molecular Biology. 1994;3(3):159-70.

Caumes E, Carrière J, Guermonprez G, Bricaire F, Danis M, Gentilini M. Dermatoses associated with travel to tropical countries: a prospective study of the diagnosis and management of 269 patients presenting to a tropical disease unit. Clinical infectious diseases. 1995;20(3):542-8.

Cepeda-Palacios R, Frugère S, Dorchies P. Expected effects of reducing *Oestrus ovis* L. mature larval weight on adult populations. Veterinary Parasitology. 2000;90(3):239-46.

Chabaudie N, Boulard C. Effect of hypodermin A, an enzyme secreted by *Hypoderma lineatum*

(Insect Oestridae), on the bovine immune system. Veterinary Immunology and Immunopathology. 1992;31(1-2):167-77.

Champer J, Yang E, Lee E, Liu J, Clark AG, Messer PW. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads through a cage population. Proceedings of the National Academy of Sciences. 2022;117(39):24377-83.

Çiftçioğlu N, Altintaş K, Haberal M. A case of human orotracheal myiasis caused by *Wohlfahrtia magnifica*. Parasitology research. 1996;83:34-6.

Concha C, Scott MJ. Sexual development in *Lucilia cuprina* (Diptera, Calliphoridae) is controlled by the transformer gene. Genetics. 2009;182:785-98.

Concha C, Palavesam A, Guerrero FD, Sagel A, Li F, Osborne JA, et al. A transgenic male-only strain of the New World screwworm for an improved control program using the sterile insect technique. BMC biology. 2016;14(1):1-13.

Concha C, Yan Y, Arp A, Quilarque E, Sagel A, de León AP, et al. An early female lethal system of the New World screwworm, *Cochliomyia hominivorax*, for biotechnology-enhanced SIT. BMC genetics. 2020;21(2):1-12.

Cruz MDS, Robles MCV, Trapman JJ, Thomas G. Comparative rearing of *Wohlfahrtia magnifica* (Diptera: Sarcophagidae) in Dead and Living tissues and the impact of cold storage on pupal survival. Journal of medical entomology. 1998;35(2):153-6.

Cruz MDS, Robles MCV, Thomas G. In vivo rearing and development of *Wohlfahrtia magnifica* (Diptera: Sarcophagidae). Journal of medical entomology. 1996;33(4):586-91.

da Silva NM, de Azeredo-Espin AM. Investigation of mutations associated with pyrethroid resistance in populations of the New World Screwworm fly, *Cochliomyia hominivorax* (Diptera: Calliphoridae). Genetics and Molecular Research. 2009;8(3):1067-78.

Diakakis N, Papadopoulos E, Hall MJR, Desiris A. Post-traumatic complication due to *Wohlfahrtia magnifica* larvae on a horse. Veterinary Record. 2006;158(5):170-2.

Elkington RA, Humphries M, Commins M, Maugeri N, Tierney T, Mahony TJ. A *Lucilia cuprina* excretory–secretory protein inhibits the early phase of lymphocyte activation and subsequent proliferation. Parasite Immunology. 2009;31(12):750-65.

Erdmann GR. Antibacterial action of myiasis-causing flies. Parasitology Today. 1987;3(7):214-6.

Farkas R, Hall MJR, Bouzagou AK, Lhor Y, Khallaayoune K. Traumatic myiasis in dogs caused by *Wohlfahrtia magnifica* and its importance in the epidemiology of wohlfahrtiosis of livestock. Medical and Veterinary Entomology. 2009;23:80-5.

Farkas R, Hall MJ, Daniel M, Börzsönyi L. Efficacy of ivermectin and moxidectin injection against larvae of *Wohlfahrtia magnifica* (Diptera: Sarcophagidae) in sheep. Parasitology research. 1996;82:82-6.

Farkas R, Hall MJR, Kelemen F. Wound myiasis of sheep in Hungary. Veterinary Parasitology. 1997;69(1-2):133-44.

Farkas R, Hell E, Hall MJR, Gyurkovszky M. In vitro rearing of the screwworm fly *Wohlfahrtia magnifica*. Medical and veterinary entomology. 2005;19(1):22-6.

Francesconi F, Lupi O. Myiasis. Clinical microbiology reviews. 2012;25(1):79-105.

Gaglio G, Brianti, E, Abbene S, Giannetto S. Genital myiasis by *Wohlfahrtia magnifica* (Diptera, Sarcophagidae) in Sicily (Italy). Parasitology Research. 2011;109:1471-4.

Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. Proceedings of the National Academy of Sciences. 2015;112(49):E6736-43.

Giangaspero A, Traversa, D, Trentini R, Scala A, Otranto D. Traumatic myiasis by *Wohlfahrtia magnifica* in Italy. Veterinary Parasitology. 2011;175(1-2):109-12.

Gong P, Epton MJ, Fu G, Scaife S, Hiscox A, Condon KC, et al. A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. Nature Biotechnology. 2005;23(4):453-6.

Gonzalez-Garay ML. Introduction to Isoform Sequencing Using Pacific Biosciences Technology (Iso-Seq). In: Wu J, editors. Transcriptomics and Gene Regulation. Dordrecht: Springer; 2016. p. 141-60.

Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(12):5547-51.

Gour S, Ramesh G, Kumar V, Thapliyal GK, Nagarajappa R. Cavitary myiasis and its management. Journal of Experimental Therapeutics and Oncology. 2018;12(3):211-6.

Hall MJR, Adams ZJO, Wyatt NP, Testa JM, Edge W, Nikolausz M, et al. Morphological and mitochondrial DNA characters for identification and phylogenetic analysis of the myiasis-causing flesh fly *Wohlfahrtia magnifica* and its relatives, with a description of *Wohlfahrtia monegrosensis* sp. n. Wyatt & Hall. Medical and veterinary entomology. 2009a;23:59-71.

Hall MJR, Testa JM, Smith L, Adams ZJO, Khallaayoune K, Sotiraki S, et al. Molecular genetic analysis of populations of Wohlfahrt's wound myiasis fly, *Wohlfahrtia magnifica*, in outbreak populations from Greece and Morocco. Medical and Veterinary Entomology. 2009b;23:72-9.

Hall MJR, Wall R. Myiasis of humans and domestic animals. Advances in Parasitology. 1995;35:257-334.

Hall MJR, Wall RL, Stevens JR. Traumatic myiasis: a neglected disease in a changing world. Annual Review of Entomology. 2016;61:159-76.

Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles*

gambiae. Nature Biotechnology. 2016;34(1):78-83.

Hope FW. On insects and their larvae occasionally found in the human body. Transactions of the Entomological Society of London. 1840;2:256.

King AM, MacRae TH. Insect heat shock proteins during stress and diapause. Annual review of entomology. 2015;60:59-75.

Kingan SB, Heaton H, Cudini J, Lambert CC, Baybayan P, Galvin BD, et al. A high-quality de novo genome assembly from a single mosquito using PacBio sequencing. Genes. 2019;10(1):62.

Knipling EF. Possibilities of insect control or eradication through the use of sexually sterile males. Journal of Economic Entomology. 1955;48:459-62.

Knipling EF. Sterile-male method of population control. Science. 1959;130:902-4.

Kotze AC, James PJ. Control of sheep flystrike: what's been tried in the past and where to from here. Australian veterinary journal. 2022;100(1-2):1-19.

Kyrou K, Hammond AM, Galizi R, Kranjc N, Burt A, Beaghton AK, et al. CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. Nature Biotechnology. 2018;36(11):1062-6.

Lehrer AZ, Lehrer M, Verstraeten C. Annales de Médecine Veterinaire. 1988;132:475.

Li F, Vensko SP, Belikoff EJ, Scott MJ. Conservation and sex-specific splicing of the transformer gene in the Calliphorids *Cochliomyia hominivorax*, *Cochliomyia macellaria* and *Lucilia sericata*. PLoS One. 2013;8(2):e56303.

Li F, Wantuch HA, Linger RJ, Belikoff EJ, Scott MJ. Transgenic sexing system for genetic control of the Australian sheep blow fly *Lucilia cuprina*. Insect Biochemistry and Molecular Biology. 2014;51:80-8.

Li F, Zhao X, Li M, He K, Huang C, Zhou Y, et al. Insect genomes: progress and challenges. Insect molecular biology. 2019a;28(6):739-58.

Li H, Oyun G, Bao H, Yunzhang L, Yang B, Liu T, et al. Morphological and scanning electron microscopic (SEM) studies of the pupae of *Wohlfahrtia magnifica*. Journal of Camel Practice and Research. 2020;27(1):17-22.

Li Y, Li X, Liu J, Liu A, Guo P, Han Y, et al. Molecular identification and detection of *Wohlfahrtia magnifica* in ovine vulvar myiasis in Gansu, China. Tropical animal health and production. 2019b;51:2629-34.

Liu J, Hou B, Wuen J, Jiang N, Gao T, Hasi S. Epidemiological Investigation on Genital Myiasis of Bactrian Camels in Parts of Inner Mongolia, China. Journal of Camel Practice and Research. 2022;29:229-35.

Marangi M, Hall MJ, Aitken A, Ready PD, Giangaspero A. Origins of *Wohlfahrtia magnifica* in Italy based on the identification of mitochondrial cytochrome b gene haplotypes. Parasitology research. 2016;115:483-7.

Martinson EO, Peyton J, Kelkar YD, Jennings EC, Benoit JB, Werren JH, et al. Genome and ontogenetic-based transcriptomic analyses of the flesh fly, *Sarcophaga bullata*. G3: Genes, Genomes, Genetics. 2019;9(5):1313-20.

Maurya RP, Mishra D, Bhushan P, Singh VP, Singh MK. Orbital myiasis: due to invasion of larvae of flesh fly (*Wohlfahrtia magnifica*) in a child; rare presentation. Case reports in Ophthalmological medicine. 2012;2012:1-2.

McGraw TA, Turiansky GW. Cutaneous myiasis. Journal of the American Academy of Dermatology. 2008;58(6):907-26.

Moiré N. Hypodermin A and inhibition of lymphocyte proliferation. Parasitology today. 1998;14(11):455-57.

Moiré N, Nicolas-Gaulard I, Le Vern Y, Boulard C. Enzymatic effect of hypodermin A, a

parasite protease, on bovine lymphocyte membrane antigens. Parasite Immunology. 1997;19(1):21-7.

Muthukrishnan S, Mun S, Noh MY, Geisbrecht ER, Arakane Y. Insect cuticular chitin contributes to form and function. Current Pharmaceutical Design. 2020;26(29):3530-45.

Noutsis C, Millikan LE. Myiasis. Dermatologic clinics. 1994;12(4):729-36.

Osorio-Pinzon J, Palencia A, Cruz-Calderon S, Rodriguez-Morales AJ. Myiasis and Tungiasis. Current Tropical Medicine Reports. 2021;8:112-20.

Parker A, Mehta K. Sterile insect technique: a model for dose optimization for improved sterile insect quality. Florida entomologist. 2007;90(1):88-95.

Pruett Jr JH. Proteolytic cleavage of bovine IgG by hypodermin A, a serine protease of *Hypoderma lineatum* (Diptera: Oestridae). Journal of Parasitology. 1993;79(6):829-33.

Remesar S, Otero JL, Panadero R, Díez-Baños P, Díaz P, García-Díos D, et al. Traumatic myiasis by *Wohlfahrtia magnifica* in sheep flocks from southeastern Spain: prevalence and risk factors. Medical and Veterinary Entomology. 2022;36(1):30-7.

Rendón P, McInnis D, Lance D, Stewart J. Medfly (Diptera: Tephritidae) genetic sexing: Large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. Journal of economic entomology. 2004;97(5);1547-53.

Robbins K, Khachemoune A. Cutaneous myiasis: a review of the common types of myiasis. International journal of dermatology. 2010;49(10):1092-98.

Ruiz Martinez I, Cruz SMD, Rodriguez R, Lopez DM, Parra MS, Navio FA. Myiasis caused by *Wohlfahrtia magnifica* in southern Spain. Israel Journal of Veterinary Medicine. 1987;43:34-41.

Ruiz Martinez I, Leclercq M. Data on distribution of screwworm fly *Wohlfahrtia magnifica* (Schiner) in southwestern Europe (Diptera: Sarcophagidae). Notes Fauniques Gembloux.

1994;28:53-60.

Ruiz Martinez I, Soler Cruz MD, Benitez Rodriguez R, Perez Jimenez JM, Diaz Lopez M. Myiasis caused by *Wohlfahrtia magnifica* in sheep and goats in southern Spain. II. Effect of age, body region and sex on larval infestation. Israel Journal of Veterinary Medicine. 1991;46:64-8.

Sandeman RM, Feehan, JP, Chandler RA, Bowles VM. Tryptic and chymotryptic proteases released by larvae of the blowfly, *Lucilia cuprina*. International Journal for Parasitology. 1990;20(8):1019-23.

Sandeman RM, Levot GW, Heath ACG, James PJ, Greeff JC, Scott MJ, et al. Control of the sheep blowfly in Australia and New Zealand–are we there yet?. International journal for parasitology. 2014;44(12):879-91.

Scholl PJ, Catts EP, Mullen GR. Myiasis (Muscoidea, Oestroidea). In: Mullen G, Durden L, editors. Medical and veterinary entomology. San Diego: Elsevier; 2019. p. 383-419.

Scott MJ, Benoit JB, Davis RJ, Bailey ST, Varga V, Martinson EO, et al. Genomic analyses of a livestock pest, the New World screwworm, find potential targets for genetic control programs. Communications Biology. 2020;3(1):424.

Scott MJ, Concha C, Welch JB, Phillips PL, Skoda SR. Review of research advances in the screwworm eradication program over the past 25 years. Entomologia Experimentalis et Applicata. 2017;164(3):226-36.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210-2.

Solomon M, Lachish T, Schwartz E. Cutaneous myiasis. Current infectious disease reports. 2016;18:1-7.

Sotiraki S, Farkas R, Hall MJR. Fleshflies in the flesh: Epidemiology, population genetics and

control of outbreaks of traumatic myiasis in the Mediterranean Basin. Veterinary Parasitology. 2010;174(1-2):12-8.

Sotiraki S, Stefanakis, A, Hall MJR. Assessment of cypermethrin and doramectin for controlling wohlfahrtiosis in Crete. Veterinary Parasitology. 2003;116(4):327-32.

Sotiraki S, Stefanakis A, Hall MJR, Graf JF. Field trial of the efficacy of dicyclanil for the prevention of wohlfahrtiosis of sheep. Veterinary record. 2005;156(2):37-40.

Tabouret G, Bret-Bennis L, Dorchies P, Jacquiet P. Serine protease activity in excretory–secretory products of *Oestrus ovis* (Diptera: Oestridae) larvae. Veterinary Parasitology. 2003;114(4):305-14.

Tandonnet S, Cardoso GA, Mariano-Martins P, Monfardini RD, Cunha VA, de Carvalho RA, et al. Molecular basis of resistance to organophosphate insecticides in the New World screw-worm fly. Parasites & Vectors. 2022;13(1):1-13.

Toth EM, Hell E, Kovács G, Borsodi AK, Marialigeti K. Bacteria isolated from the different developmental stages and larval organs of the obligate parasitic fly, *Wohlfahrtia magnifica* (Diptera: Sarcophagidae). Microbial ecology. 2006;51:13-21.

Toth EM, Schumann P, Borsodi AK, Keki Z, Kovacs AL, Marialigeti K. *Wohlfahrtiimonas chitiniclastica* gen. nov., sp. nov., a new gammaproteobacterium isolated from *Wohlfahrtia magnifica* (Diptera: Sarcophagidae). International Journal of Systematic and Evolutionary Microbiology. 2008;58(4):976-81.

Tóth E, Kovács G, Schumann P, Kovács AL, Steiner U, Halbritter A, et al. Schineria larvae gen. nov., sp. nov., isolated from the 1st and 2nd larval stages of *Wohlfahrtia magnifica* (Diptera: Sarcophagidae). International Journal of Systematic and Evolutionary Microbiology. 2001;51(2):401-7.

Valentin A, Baumann MPO, Schein E, Bajanbileg S. Genital myiasis (Wohlfahrtiosis) in camel herds of Mongolia. Veterinary Parasitology. 1997;73(3-4):335-46.

Wyss JH. Screwworm eradication in the Americas. Annals of the New York Academy of Sciences. 2000;916:186-93.

Yadav AK, Butler C, Yamamoto A, Patil AA, Lloyd AL, Scott MJ. CRISPR/Cas9-based split homing gene drive targeting doublesex for population suppression of the global fruit pest *Drosophila suzukii*. Proceedings of the National Academy of Sciences. 2023;120(25):e2301525120.

Yan L, Zhang M, Tang L, Ente M, Ma X, Chu H, et al. First reports of nasal and traumatic myiasis infection in endangered Przewalski's horses (*Equus ferus przewalskii*). International Journal for Parasitology: Parasites and Wildlife. 2019;9:21-4.

Yan Y, Scott MJ. Building a transgenic sexing strain for genetic control of the Australian sheep blow fly *Lucilia cuprina* using two lethal effectors. BMC genetics. 2020;21:1-11.

Yasuda M. Morphology of the Larva of *Wohlfahrtia magnifica* Schin. Found in a Wound on a Camel in Inner Mongolia. Journal of Chosen Natural History Society. 1940;7(29):27-36.

Yi HY, Chowdhury M, Huang YD, Yu XQ. Insect antimicrobial peptides and their applications. Applied microbiology and biotechnology. 2014;98:5807-22.

Zhang D, Yan L, Zhang M, Chu H, Cao J, Li K, et al. Phylogenetic inference of calyptrates, with the first mitogenomes for Gasterophilinae (Diptera: Oestridae) and Paramacronychiinae (Diptera: Sarcophagidae). International Journal of Biological Sciences. 2016;12(5):489.

Zumpt F. Myiasis in man and animals in the Old World. London: Butterworths; 1965.