Genetics and ecology of an unusual sex ratio distorter in the booklouse *Liposcelis* sp.

by

Caitlin I. Curtis
B.Sc., University of Victoria, 2015

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biology

Supervisory Committee

Genetics and ecology of an unusual sex ratio distorter in the booklouse *Liposcelis* sp.

by

Caitlin I. Curtis
B.Sc., University of Victoria, 2015

Supervisory Committee

Dr. Steve J. Perlman, Department of Biology
Supervisor

Dr. Jürgen Ehlting, Department of Biology
Departmental Member

Dr. Christopher J. Nelson, Department of Biochemistry and Microbiology
Outside Member

Abstract

Supervisory Committee
Dr. Steve J. Perlman, Department of Biology
Supervisor
Dr. Jürgen Ehlting, Department of Biology
Departmental Member
Dr. Christopher J. Nelson, Department of Biochemistry and Microbiology
Outside Member

Selfish genetic elements can distort the sex ratios of their hosts by increasing their own transmission to the next generation in a non-mendelian fashion. These elements can be either nuclear genes on a sex chromosome or cytoplasmically inherited microbes, and achieve an increased transmission by manipulating gametogenesis or host reproduction. Often these selfish elements benefit from a female biased population (for example heritable microbes are passed on maternally in the egg cytoplasm), while non-selfish, autosomal genes are selected to produce a balanced sex ratio. These differing reproductive strategies cause a genetic conflict that results in an "evolutionary arms race" that can promote the evolutionary change of sex determination systems. In this thesis, I investigate an extreme sex ratio distortion in a species of booklouse, *Liposcelis* sp. This species contains two distinct female types, one of which carries a maternally transmitted selfish genetic element that results in exclusively female offspring being produced. Recently, a candidate for the sex ratio distortion was identified as a horizontally transferred bacterial gene, that we have called Odile, and that is present in the genome of the (distorter) female carrying the distorting element. The gene originates from the endosymbiotic bacterium *Wolbachia* that is well known for its ability to distort the sex ratio of its hosts.

I investigated this horizontal gene transfer event and attempt to characterize Odile. I provide evidence that this *Wolbachia* gene has been integrated into the genome of the distorter females and is not a bacterial contaminant. I found that the Odile gene has been duplicated and may have been horizontally transferred from *Wolbachia* independently to at least three other insect genomes. Additionally, I found that Odile is

transcribed at low levels in a life-stage specific manner that is suggestive of a role in development. Additionally, I looked into male mate choice in this species as one aspect of the persistence of the distorting element. I found that male *Liposcelis* sp. do not discriminate between the two female types and do not spend more time mating with one female type over the other. These results contribute to ongoing research into the extreme sex ratio distortion found in this species and the candidate gene that may be the cause. Selfish genetic elements are an important driver of sex determination evolution, and *Liposcelis* sp. provides a unique and exciting system to investigate the implications of selfish elements in a genome further.

Table of Contents

List of Tables

List of Figures

Acknowledgments

There are so many people I would like to acknowledge who have provided me with advice and support over the course of my MSc degree. I would first like to thank my supervisor Steve Perlman for all of his support and guidance. Steve is a big part of why the time in his lab has been such a positive experience. I would also like to thank my committee members, Chris Nelson and Jürgen Ehlting for their advice and guidance, as well as my first mentor Patrick Walter. Thank you to Matthew Ballinger, Ryan Gawryluk, Graeme Keais and Haewon Shin for answering my many questions and making me laugh every day. I couldn't have asked for a better group of lab members to share this experience with. I would also like to give a special thank you to Sarah Ewing for being there with me every step of the way. Finally, thank you to my parents and to Mike who have supported me immeasurably.

Chapter 1 – Selfish genetic elements as sex ratio distorters

The sex ratio of males to females was classically described to be at an equilibrium when there was a 1:1 ratio of male to female offspring (Fisher, 1930). Fisher explained this equilibrium in terms of 'parental expenditure', in which a ratio of 1:1 is favoured if the effort to produce the two sexes is equal. Fisher's principle has been summarised and restated by Bodmer and Edwards (1960) in order to explain the Fisherian sex ratio without relating it to parental expenditure. In this case, if the sex ratio deviates from 1:1, the rarer sex would have a higher number of mating prospects, and parents genetically disposed to produce more of the rarer sex will therefore have a higher fitness. As more of the rarer sex are produced and the 1:1 sex ratio is again approached, the advantage of producing the rarer sex dissipates. Although Fisher's theory can apply to many species, it does not hold true to all. There are many examples of species that have sex ratios that deviate from the classic 1:1 ratio of males to females (Hamilton, 1967; Werren, 1987), and these deviations can occur for a number of reasons, including conflicts over transmission. Some genetic elements are transmitted to offspring in a non-mendelian, selfish manner, and if these elements influence sexual reproduction, sex ratio biases can occur (Doolittle and Sapienza, 1980; Orgel and Crick, 1980; Werren et al. 1988).

Selfish genetic elements gained their name because of the 'selfish' nature in which they replicate and are transmitted from generation to generation. These genetic elements have gained a transmission advantage despite being either neutral or detrimental to the organism as a whole (Doolittle and Sapienza, 1980; Orgel and Crick, 1980; Werren et al. 1988). Genetic elements that increase their transmission relative to that of related genes are now known to be common features of eukaryotes. These selfish genetic elements exist in many forms, such as transposable elements, nuclear genes, maternally inherited organelles or heritable microorganisms (Hurst and Werren, 2001). Because of the selfish nature of these elements, intragenomic conflict can arise as a response to selfish elements with differing transmission strategies (Cosmides and Tooby, 1981). It is this genetic conflict between selfish genetic elements and the resulting genes that oppose them that can drive the evolution of genome structure, new genes, new species and new

mechanisms of sex determination (Burt and Trivers, 2006). My thesis will focus specifically on selfish genetic elements that cause sex ratio distortions in their hosts.

Nuclear sex ratio distorters

Selfish genetic elements that are nuclear encoded distort the sex ratios of their hosts through a process termed meiotic drive (Sandler and Novitski, 1957). These genes, 'meiotic drivers', are transferred to the next generation in excess of the expected Mendelian ratio of 50%. This manipulation of gametogenesis is achieved in two main ways: by causing preferential segregation to the ova in females, or by the elimination of sperm that do not carry the meiotic drivers (Werren, 2011). In the first case, chromosomal variants can bias their own transmission to the maturing egg cell over their non-driving homologs. For example, nuclear elements such as B chromosomes, (non-essential chromosomes), are widely found in plants and animals and can accumulate selfishly, in some cases preferentially segregating to the ovum of females and/or manipulate the sex ratios of their hosts (Camacho et al. 2000). In an extreme case found in the haplodiploid wasp *Nasonia vitripennis*, a B chromosome called paternal sex ratio (PSR), destroys the paternal chromosomes it is transmitted with (Beukeboom, 1994; Nur et al. 1988; Werren and Stouthamer, 2003). Because of the nature of haplodiploid sex determination, diploid females then develop into haploid males.

In males, meiotic drivers can increase their transmission by killing sperm that do not carry the driving element. Sex ratio distortions can result when this meiotic driver is located on a sex chromosome (Hamilton, 1967). Genes on a driving sex chromosome are disproportionately transmitted to the next generation (more than 50%). Although Y-chromosome drive exists, X-chromosomes make up the vast majority of driving sex chromosomes and have been identified in a wide range of plants, animals and insects (Jaenike, 2001). Driving X-chromosomes increase their transmission to the next generation by eliminating or disrupting Y-bearing sperm during spermatogenesis (Jaenike, 2001). The result is a distorted sex ratio in favour of females, often at very high frequencies. For example, in the fly *Drosophila pseudoobscura*, the driving X-

chromosome called SR (Sex Ratio), causes the failure of all Y-bearing sperm and results in offspring that are 100% female (Policansky and Ellison, 1970).

Cytoplasmic sex ratio distorters

Cytoplasmic elements such as organelles (mitochondria and chloroplasts), and endosymbiotic microbes are passed on to the next generation via egg cytoplasm, with a few exceptions (Cosmides and Tooby, 1981). These cytoplasmically inherited genetic elements therefore benefit from a female biased sex ratio as they are transmitted almost solely by females and not by males. However, nuclear genes have their own transmission strategies that often favour an equal investment in male and female offspring (Fisher, 1930). These two competing transmission strategies result in a genetic conflict and have thus likely played a role in sex-determination evolution (Werren, 2011).

There are many examples of microbes that have exploited this inheritance mechanism in order to increase the number of female offspring and therefore their own transmission. Host biology has been manipulated to produce more females by way of inducing parthenogenesis; feminization of males, in which infected males are converted to functional females; and by male-killing, where male embryos die while female embryos live (Engelstädter and Hurst, 2009). The endosymbiotic bacterium *Wolbachia* has been the best studied microbial manipulator of sex ratios as it is widespread in insects and can manipulate host reproduction in all of these ways (Stouthamer et al. 1999; Werren et al. 2008). *Wolbachia*-induced feminization was first reported in the pillbug *Armadillidium vulgare*, where *Wolbachia*-infected males develop as functional phenotypic females (Rigaud et al. 1991). In the butterfly *Acraea encodon*, Jiggins et al. (2002) reported a case of male-killing in which 95% of females are infected with *Wolbachia* and consequently produce only daughters. Similarly, an extreme sex ratio distortion is found in the Polynesian butterfly *Hypolimnus bolina*, where 99% of the females are infected with *Wolbachia* and produce 100 females for every one male (Dyson and Hurst, 2004).

Mechanism of sex ratio distortion

The specific mechanisms used by both nuclear and cytoplasmic selfish genetic elements to distort sex ratios are mostly unknown. There are however a few cases in which the mechanisms have been described. Previous cytological studies of X-drive in *Drosophila* species have shown irregular Y-chromosome behaviour in meiosis II leading to non-functional sperm (Jaenike, 2001). A study by Helleu et al. (2016) demonstrated that the X-drive in *Drosophila simulans* was caused by variant of a heterochromatin protein (HP1D2), expressed in the male germline that specifically binds the Y-chromosome and thus prevents segregation during meiosis II. Recently, Aldrich et al. (2017) identified a mechanism involved in the elimination of paternal chromosomes by the selfish B-chromosome (PSR) in the wasp *Nasonia vitripennis*. Three specific histone marks were disrupted by PSR, preventing chromosome formation during the first embryonic mitosis. Interestingly, PSR was missing two of the histone marks allowing PSR to avoid self-elimination.

As with nuclear selfish genetic elements, there are few cases in which the mechanism of distortion by heritable microbes has been explained. Although the exact molecular mechanisms of *Wolbachia*-induced feminization are not known, the bacterium may be acting on the 'male' gene which controls the development of the androgenic gland in the pillbug *A. vulgare* (Cordaux et al. 2011). It was found that androgenic gland function is disrupted in *Wolbachia*-infected males preventing the differentiation of primary and secondary male characteristics (Badawi et al. 2015; Martin et al. 1999; Suzuki & Yamasaki, 1997). Recently, a gene from the bacterium *Spiroplasma poulsonii* was identified as the potential cause of male-killing in *Drosophila melanogaster* (Harumoto and Lemaitre, 2018). Overexpression of the gene *Spaid*, kills males (but not females) by inducing apoptosis and neural defects. Microbial induced parthenogenesis has thus far only been identified in haplodiploid species in which males develop from unfertilized eggs. The specific mechanisms behind the induction of parthenogenesis are unclear; however, embryonic development is manipulated to produce unfertilized diploid eggs that develop into female offspring (Werren et al. 2008).

Persistence of sex ratio distorters

Selfish genetic elements manipulate host reproduction in order to increase their transmission above the expected Mendelian ratio. Models of population genetics have determined three important factors involved in the spread of selfish sex ratio distorters: the penetrance of the induced phenotype; the transmission rate to offspring; and the symbiont's effect on female fitness (Engelstädter and Hurst, 2009). If a selfish sex ratio distorter spreads within a host population leading to a critical deficiency of one sex, it would eventually induce extinction of the host population and itself unless the species evolves a parthenogenetic mode of sex determination (Hatcher et al. 1999). Why then do we see many examples of selfish genetic elements that haven't spread to fixation and thereby resulting in the extinction of their hosts? There is a large assortment of factors that allow these selfish elements to persist over time, such as population structure and population size, intragenomic conflict, and mating habits (Hatcher 2000; Hurst and Werren 2001).

The presence of selfish genetic elements creates a genetic conflict between the selfish element and the rest of the host genome. Because of this, genes that suppress or resist these selfish elements are selected for (Hurst et al. 1996). In cases of X-chromosome drive, autosomal genes can often evolve to suppress the action of the driving X. In a female biased population, these genes will be selected for as the rarer sex, in this case males, have an increase in fitness (Fisher, 1930). Additionally, suppressors evolve on the Y-chromosome in order to avoid the direct action of the driving X (Jaenike, 2001). Suppressors also evolve in response to sex ratio distorting microbes. An interesting case of suppression was discovered in an *H. bolina* butterfly population in Southeast Asia that is infected with the same strain of *Wolbachia* that causes male-killing at extremely high levels in the Polynesian population of *H. bolina* (Hornett et al. 2006). The Southeast Asian population suppresses the male-killing phenotype of *Wolbachia* and produces a 1:1 sex ratio of males to females (compared to the 1:100 ratio in the Polynesian population). Through introgression with the Polynesian strain, it was demonstrated that males are rescued by a dominant suppressor of male-killing in the host genome. Two nuclear suppressors have also evolved against the feminizing action of *Wolbachia* in the pillbug *A. vulgare*. In the first case, the suppressor overrides the

feminizing effect and infected males will develop as males. In the second case, the suppressor prevents the transmission of the feminizing *Wolbachia* to the next generation but does not override the feminizing effect and males will develop as females (Caubet et al. 2000).

In addition to genetic suppressors that evolve in response to selfish genetic elements, mating preferences can arise against individuals that harbour a sex ratio distorter. The discrimination against individuals that carry these selfish elements can act to control the spread in a population and prevent the selfish element from reaching fixation. Alternatively, if a mate does not discriminate against a sex ratio distorting element, this could allow for the element to persist in the population without being lost. Mate choice against a selfish X-chromosome has been demonstrated in the stalk-eyed fly (*Teleopsis dalmanni* and *T. whitei*), in which females preferentially mate with males that do not carry the driving X, and that can be distinguished by a shorter eye-stalk length (Wilkinson et al. 1998). As well, in the pillbug *A. vulgare*, males prefer to mate with uninfected females compared to *Wolbachia*-infected feminized males (Moreau et al. 2001).

Sex ratio distortion in the booklouse Liposcelis sp.

An extreme sex ratio distortion was recently discovered in a yet unnamed species of booklice, *Liposcelis* sp., originally collected in the Chiricahua Mountains, Arizona in 2010 (Perlman et al. 2015). Booklice belong to the Order Psocodea, that also includes barklice and parasitic lice, and are often of interest because of their agricultural status as a stored grain pest. This species of booklouse is sexual and contains two distinct female types. The sex ratio distortion in this species is caused by a selfish genetic element that is carried by one of the female types, termed distorter females (Perlman et al. 2015). Distorter females produce exclusively female offspring, while the other female type, termed nondistorter females, produce a mixed sex ratio of both males and females. Both female types must mate with a male in order to reproduce; in other words, distorter females mate with the sons of nondistorter females. It was determined that the selfish genetic element, termed distorting element, was not a heritable microbe but a nuclear element (Perlman et al. 2015). Additionally, it is maternally transmitted from the females

that carry it to all of their offspring, which in turn only produce female offspring. Distorter females are diploid, but only ever transmit the genes from their mothers to the next generation and new paternal genes are introduced each generation (Hamilton et al. 2018). Because of this, the distorter-restricted genomes do not recombine in the population and are diverging from the nondistorter genomes.

Recently, a candidate gene for sex ratio distortion in *Liposcelis* sp. was identified by comparing the distorter and nondistorter female genomes (Hamilton et al. 2018). The candidate appeared to be horizontally transferred from the bacterium *Wolbachia* and was only present in the distorter genome. This gene was of particular interest as *Wolbachia* is a well-known reproductive manipulator. In this thesis, I investigate this gene, which we have called Odile, as the candidate for the sex ratio distortion in *Liposcelis* sp. My work is divided into two chapters. In chapter 2, I characterize Odile with the use of bioinformatics and experimental techniques. I present evidence that the Odile gene is integrated in the distorter genome and is not bacterial contamination. I looked into the evolutionary relationship of Odile, and investigated the presence of conserved domains and other biological properties. Additionally, I investigated whether Odile shows a life-stage specific expression pattern that may be suggestive of function. In chapter 3, I looked into one of the factors that may have allowed this selfish genetic element to persist over time, specifically, the effect of male mate choice. I wanted to investigate whether male *Liposcelis* sp. discriminate between distorter and nondistorter females in mate choice trials. I also looked into whether there were significant differences between the two female types in the duration of mating and the time it took for mating to initiate.

This research provides an initial characterization of a horizontal gene transfer event from a symbiotic bacterium to an insect genome that may be causing an extreme sex ratio distortion in a species of booklouse. *Liposcelis* sp. is an exciting system to study as booklice and parasitic lice have an unusual baseline form of reproduction, called paternal genome elimination (PGE) (Hodson et al. 2017). In this mode of sex determination, males develop from fertilized eggs, but the paternal chromosomes are either inactivated or eliminated in the embryos that are destined to become males (Brown and Nur 1964; Sanchez, 2008). As the distorter *Liposcelis* sp. do not produce males, it appears as if PGE has been hijacked by the distorters, which results in only females being

produced. In addition, distorter females in this species carry a selfish genetic element that is maternally transmitted and causes an extreme sex ratio distortion in the species. This unique system provides an exciting opportunity to investigate a chromosomal sex ratio distorter that is proposed to have originated from a bacterium well known to distort the sex ratios of its host.

Chapter 1 – References

Aldrich, J.C., Le, A., Cheema, M.S., Ausió, J., and Ferree, P.M. 2017. A 'selfish' B chromosome induces genome elimination by disrupting the histone code in the jewel wasp *Nasonia vitripennis*. Sci. Rep. **7**(1). doi:10.1038/srep42551.

Badawi, M., Grève, P., and Cordaux, R. 2015. Feminization of the isopod Cylisticus convexus after transinfection of the wVulC *Wolbachia* strain of *Armadillidium vulgare*. PLoS One **10**(6). doi:10.1371/journal.pone.0128660.

Beukeboom, L.W. 1994. Phenotypic fitness effects of the selfish B chromosome, paternal sex ratio (PSR) in the parasitic wasp *Nasonia vitripennis*. Evol. Ecol. **8**(1): 1–24. doi:10.1007/BF01237662.

Bodmer, W.F., and Edwards, A.W. 1960. Natural selection and the sex ratio. Ann. Hum. Genet. **24**(3): 239–244.

Burt, A., and Trivers, R. 2006. Genes in Conflict: The Biology of Selfish Genetic Elements. Belknap Press, Cambridge, MA.

Camacho, J.P.M., Sharbel, T.F., and Beukeboom, L.W. 2000. B-chromosome evolution. Philos. Trans. R. Soc. London B Biol. Sci. **355**(1394): 163–178. doi:10.1098/rstb.2000.0556.

Caubet, Y., Hatcher, M.J., Mocquard, J.P., and T. Rigaud. 2000. Genetic conflict and changes in heterogametic mechanisms of sex determination. J. Evol. Biol. **13**(5): 766–777. doi:10.1046/j.1420-9101.2000.00225.x.

Cordaux, R., Bouchon, D., and Grève, P. 2011. The impact of endosymbionts on the evolution of host sex-determination mechanisms. Trends Genet. **27**(8): 332–341. doi:10.1016/j.tig.2011.05.002.

Cosmides, L.M., and Tooby, J. 1981. Cytoplasmic inheritance and intragenomic conflict. J. Theor. Biol. **89**(1): 83–129.

Doolittle, W.F., and Sapienza, C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature **284**(5757): 601–603. doi:10.1038/284601a0.

Dyson, E.A., and Hurst, G.D.D. 2004. Persistence of an extreme sex-ratio bias in a natural population. Proc. Natl. Acad. Sci. **101**(17): 6520–6523. doi:10.1073/pnas.0304068101.

Engelstädter, J., and Hurst, G.D.D. 2009. The ecology and evolution of microbes that manipulate host reproduction. Annu. Rev. Ecol. Evol. Syst. **40**: 127–149. doi:10.1146/annurev.ecolsys.l.

Fisher, R.A. 1930. The Genetical Theory Of Natural Selection. Clarendon Press, Oxford, England.

Hamilton, P.T., Hodson, C.N., Curtis, C.I., and Perlman, S.J. 2018. Genetics and genomics of an unusual sex ratio distortion in an insect. Curr. Biol. **28**(23): 3864-3870.

Hamilton, W.D. 1967. Extraordinary sex ratios. Science **156**(3774): 477–488. doi:10.1126/science.156.3774.477.

Harumoto, T., and Lemaitre, B. 2018. Male-killing toxin in a bacterial symbiont of *Drosophila*. Nature **557**(7704): 252–255. doi:10.1038/s41586-018-0086-2.

Hatcher, M.J. 2000. Persistence of selfish genetic elements: Population structure and conflict. Trends Ecol. Evol. **15**(7): 271–277. doi:10.1016/S0169-5347(00)01875-9.

Hatcher, M.J., Taneyhill, D.E., Dunn, A.M., and Tofts, C. 1999. Population dynamics under parasitic sex ratio distortion. Theor. Popul. Biol. **56**(1): 11–28. doi:10.1006/tpbi.1998.1410.

Helleu, Q., Gérard, P.R., Dubruille, R., Ogereau, D., Prud'homme, B., Loppin, B., and Montchamp-Moreau, C. 2016. Rapid evolution of a Y-chromosome heterochromatin protein underlies sex chromosome meiotic drive. Proc. Natl. Acad. Sci. **113**(15): 4110–4115. doi:10.1073/pnas.1519332113.

Hodson, C.N., Hamilton, P.T., Dilworth, D., Nelson, C.J., Curtis, C.I., and Perlman, S.J. 2017. Paternal Genome Elimination in *Liposcelis* Booklice. **206**(June): 1091–1100. doi:10.1534/genetics.117.199786/-/DC1.1.

Hornett, E.A., Charlat, S., Duplouy, A.M.R., Davies, N., Roderick, G.K., Wedell, N., and Hurst, G.D.D. 2006. Evolution of male-killer suppression in a natural population. PLoS Biol. **4**(9): 1643–1648. doi:10.1371/journal.pbio.0040283.

Hurst, G.D.D., and Werren, J.H. 2001. The role of selfish genetic elements in eukaryote evolution. Nat. Rev. Genet. **2**(8): 597–606.

Hurst, L.D., Atlan, A., and Bengtsson, B.O. 1996. Genetic conflicts. Q. Rev. Biol. **71**(3): 317–364.

Jaenike, J. 2001. Sex chromosome meiotic drive. Annu. Rev. Ecol. Seystematics **32**(2001): 25–49.

Jiggins, F.M., Randerson, J.P., Hurst, G.D.D., and Majerus, M.E.N. 2002. How can sex ratio distorters reach extreme prevalences? Male-killing *Wolbachia* are not suppressed and have near-perfect vertical transmission efficiency in *Acraea encedon*. Evolution (N. Y). **56**(11): 2290–2295.

Martin, G., Sorokine, O., Moniatte, M., Bulet, P., Hetru, C., and Van Dorsselaer, A. 1999. The structure of a glycosylated protein hormone responsible for sex determination in the isopod, *Armadillidium vulgare*. doi:10.1046/j.1432-1327.1999.00442.x.

Moreau, J., Bertin, A., Caubet, Y., and Rigaud, T. 2001. Sexual selection in an isopod with *Wolbachia*-induced sex reversal: Males prefer real females. J. Evol. Biol. **14**(3): 388–394. doi:10.1046/j.1420-9101.2001.00292.x.

Nur, U., Werren, J.H., Eickbush, D.G., Burke, W.D., and Eickbush, T.H. 1988. A "selfish" B chromosome that enhances its transmission by eliminating the paternal genome. Science. **240**(4851): 512–514. doi:10.1126/science.3358129.

Orgel, L.E., and Crick, F.H.C. 1980. Selfish DNA: the ultimate parasite. Nature **284**(5792): 645–646. doi:10.1038/288645a0.

Perlman, S.J., Hodson, C.N., Hamilton, P.T., Opit, G.P., and Gowen, B.E. 2015. Maternal transmission, sex ratio distortion, and mitochondria. PNAS. **112**(33): 10162–10168. doi:10.1073/pnas.1421391112.

Policansky, D., and Ellison, J. 1970. "Sex Ratio" in *Drosophila pseudoobscura*: Spermiogenic failure. Science. **169**(3948): 888–889.

Rigaud, T., Juchault, P., and Mocquard, J.P. 1991. Experimental study of temperature effects on the sex ratio of broods in terrestrial Crustacea *Armadillidium vulgare* Latr. Possible implications in natural populations. J. Evol. Biol. **4**(4): 603–617. doi:10.1046/j.1420-9101.1991.4040603.x.

Sandler, L., and Novitski, E. 1957. Meiotic drive as an evolutionary force. Am. Nat. **91**(857): 105–110.

Stouthamer, R., Breeuwer, J.A.J., and Hurst, G.D.D. 1999. *Wolbachia Pipientis*: Microbial manipulator of arthropod reproduction. Annu. Rev. Microbiol. **53**(1): 71–102. doi:10.1146/annurev.micro.53.1.71.

Suzuki, S., and Yamasaki, K. 1997. Sexual bipotentiality of developing ovaries in the terrestrial isopod *Armadillidium vulgare* (malacostraca, crustacea). Gen. Comp. Endocrinol. **107**(1): 136–146. doi:10.1006/gcen.1997.6914.

Werren, J.H. 1987. Labile sex ratios in wasps and bees: life history influences the ratio of male and female offspring. Bioscience **37**(7): 498–506.

Werren, J.H. 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc. Natl. Acad. Sci. U. S. A. **108**(2): 10863–10870. doi:10.1073/pnas.

Werren, J.H., Baldo, L., and Clark, M.E. 2008. *Wolbachia*: Master manipulators of invertebrate biology. Nat. Rev. Microbiol. **6**(10): 741–751. doi:10.1038/nrmicro1969.

Werren, J.H., Nur, U., and Wu, C.I. 1988. Selfish genetic elements. Trends Ecol. Evol. **3**(11): 297–302. doi:10.1016/0169-5347(88)90105-X.

Werren, J.H., and Stouthamer, R. 2003. PSR (paternal sex ratio) chromosomes: The ultimate selfish genetic elements. Genetica **117**(1): 85–101. doi:10.1023/A:1022368700752.

Wilkinson, G.S., Kahler, H., and Baker, R.H. 1998. Evolution of female mating preferences in stalk-eyed flies. Behav. Ecol. **9**(5): 525–533.

Chapter 2 – Characterization of a horizontally transferred bacterial gene as a candidate for sex ratio distortion in *Liposcelis* sp.

Introduction

The large majority of sexually reproducing species produce approximately a 1:1 sex ratio of male to female offspring. Fisher's principle states that if the population deviates from equal numbers of males to females, the rarer sex will gain an advantage and the population will eventually return to a stable equilibrium of 1:1 (Fisher, 1930). There are however many exceptions to this principle in which offspring sex ratios are skewed away from the 1:1 ratio of males to females and have been referred to as adaptive (Boomsma, 1991). For example, in the eusocial bee (*Augochlorella striata*), the sex ratio is predicted to be biased towards females to maximize their inclusive fitness, as workers are more closely related to their sisters than their brothers (Mueller, 1991). Similar deviations in sex ratios are seen in wasps, ants and other insects in the hymenopteran order which have a haplodiploid system of sex determination in which haploid males develop from unfertilized eggs and diploid females develop from fertilized eggs (Heimpel and de Boer, 2008).

Deviations from the predicted Fisherian sex-ratio are not always considered adaptive however, and can often occur as a result of selfish genetic elements that are able to increase their own transmission to the next generation, despite the cost to the individual's fitness (Burt and Trivers, 2006). Genetic conflict can occur as a result of differing transmission strategies and can result in sex ratio distortions if the conflict is occurring over a reproductive strategy (Werren, 2011). Examples of selfish genetic elements include regions of the nuclear genome, for instance transposable elements or driving X chromosomes, as well as cytoplasmic elements such as organelles (mitochondria and chloroplasts) or heritable microbes (Doolittle and Sapienza, 1980; Werren et al. 1988).

Nuclear selfish elements have been found to distort sex ratios in a number of ways. For example, the biased transmission of driving X chromosomes often results in

offspring that are 100% female following the prevention of functional Y-bearing sperm (Jaenike, 2001). A selfish, non-essential B chromosome in the wasp *Nasonia vitripennis* completely eliminates paternal chromosomes from being transmitted resulting in only male offspring because of the nature of the haplodiploid sex determination (Nur et al. 1988).

Endosymbiotic bacteria are a well-studied example of sex ratio-distorting cytoplasmic elements. Maternally inherited bacteria are transmitted to offspring via egg cytoplasm, thus benefitting from a female bias in the population as they are not transmitted through males (Hurst and Werren, 2001). By manipulating the sex ratio of their hosts, these microbes can increase their own transmission to future generations. This disruption of host biology by bacteria has been achieved through male-killing, cytoplasmic incompatibility, feminization of males and the induction of parthenogenesis (Engelstädter and Hurst, 2009). Although there are a number of bacterial symbionts that manipulate their host sex ratios, *Wolbachia*, a widespread endosymbiont of insects, has the reputation of being the 'master manipulator' as it can manipulate its host reproduction in each of the aforementioned ways (Engelstädter and Hurst, 2009; Hunter et al. 2003; Werren, 1997, 2011).

It has become evident that these cytoplasmic elements can have an additional effect on host evolution through the horizontal transfer of their genes to the genomes of their hosts. Recently, a 3 MB insert of a *Wolbachia* genome was discovered to be integrated into the genome of the pillbug *Armadillidium vulgare* (Leclercq et al. 2016). The insert is derived from a strain of endosymbiotic *Wolbachia* that was previously found to turn male *A. vulgare* into females. This recent transfer event resulted in a female sex-determining region that is now acting as a new sex chromosome in the pillbug. In this case, the horizontal transfer from the feminizing *Wolbachia* resulted in a change in the sex determination system of this insect. Horizontal gene transfers (HGTs) from bacteria to other bacteria are common, but less common are transfers from bacteria to multicellular eukaryotes (Beiko et al. 2005). Genetic transfers of this nature, from bacteria to multicellular eukaryotes, may be facilitated by the close proximity of bacterial endosymbionts and their hosts (Dunning Hotopp, 2011). More specifically, DNA from endosymbionts that live within eukaryotic germlines has a greater potential to be

transferred and passed on to host offspring and future generations. Horizontal gene transfers can therefore provide a mechanism for the acquisition of new genes and potentially new function as was seen in *A. vulgare*.

Although there are multiple reports of HGT between different bacteria and their hosts, the majority of cases involve the transfer of genetic material from the endosymbiotic bacteria *Wolbachia*. It has even been suggested that horizontal gene transfers from *Wolbachia* to the genome of their hosts have occurred in 70% of *Wolbachia*-infected hosts (Dunning Hotopp, 2011). The higher prevalence of HGT events involving *Wolbachia* is likely because of its widespread infection of insect species, with recent estimates predicting around 40% - 66% of insect species to be infected (De Oliveira et al. 2015; Hilgenboecker et al. 2008; Zug and Hammerstein, 2012). Almost always, DNA transferred from *Wolbachia* to host genomes result in non-functional, pseudogenized genes that contain stop codons or frameshift mutations (Dunning-Hotopp et al. 2007). Less frequently, horizontally transferred genes of *Wolbachia* origin are transcribed as is seen in *Drosophila ananassae* from Hawaii (Dunning-Hotopp et al. 2007). In this case, nearly an entire *Wolbachia* genome was integrated into the fly nuclear genome, but of the 1206 genes assayed, only 2% were transcribed. To date, there has not been a report of a horizontally transferred *Wolbachia* gene with a demonstrated function.

An extreme sex ratio distortion was recently discovered in a species of booklouse, *Liposcelis* sp., in which a selfish genetic element is manipulating the reproduction of its host (Perlman et al. 2015). There are two female types present in this species, one of which, the distorter female, carries a selfish genetic element (called the distorting element), resulting in the production of only daughters. The second female type, the nondistorter female, produces a mixed sex ratio of male and female offspring. The distorting element is transmitted maternally, as the daughters of distorters will also only produce female offspring. Distorter females are not parthenogenetic however, as they must mate with the male offspring of the nondistorter *Liposcelis* sp. in order to reproduce. Additionally, the distorting element is predicted to be a nuclear entity as there was no evidence of microbial symbionts following an extensive genomic and microscopic search (Perlman et al. 2015). Interestingly, the paternal genes are expressed within the somatic tissue of distorter offspring (see Chapter 3), but are not passed on to the next generation.

Following the genome comparison of the nondistorter and distorter females, an exciting candidate for the selfish element causing the sex ratio distortion surfaced. A gene specific to the distorter genome had *Wolbachia* origin and appeared to be integrated in the genome of the distorters as a result of a past HGT event. We named this gene Odile (Hamilton et al. 2018), for '*O*nly *D*aughters in *Liposcelis*-associated *E*lement'. Odile is the name of the Black Swan in Tchaikovsky's ballet, Swan Lake, in which Odile attempts to steal the prince from her sister, the White Swan.

I utilized a number of molecular techniques and bioinformatics to characterize this gene further. I wanted to confirm genomic integration and investigate whether the putative HGT event was a viable candidate for the sex ratio distortion in *Liposcelis* sp. In order to look into potential protein function, I looked into the presence of conserved domains, motifs and other gene features. As well, I wanted to look into the phylogenetic history of the putative HGT and the relationship to the *Wolbachia* homologs. Finally, I looked into life stage specific gene expression as both an additional confirmation of integration and as a way to provide more information on the predicted gene function. As previously shown, horizontal gene acquisition by eukaryotes can lead to novel functions or even result in a new sex chromosome (Leclercq et al. 2016). This putative HGT event from *Wolbachia* to the genome of *Liposcelis* sp. distorters presents an exciting opportunity to characterize a candidate for the selfish genetic element causing an extreme female-biased sex ratio distortion.

Methods

Booklouse colony information

The booklice used in the work for this chapter are maintained in separate (i.e. distorter and nondistorter) colonies in the laboratory and are kept at 27˚C and 75% relative humidity using a saturated NaCl solution. The distorter and nondistorter *Liposcelis* sp. were collected from the Chiricahua Mountains in southeastern Arizona in 2010. Cultures are kept in 125ml glass canning jars that contain a 1:10 (by weight) mixture of Rice Krispies (Kellogg's) to cracked wheat (Bob's Red Mill). The lids of the

jars are replaced with a 70mm Whatman filter paper (Sigma-Aldrich). Every two weeks, I maintain the colonies by replacing approximately half of the colony with fresh food and adding males from the nondistorter colonies to the distorter colonies.

Cultures of the asexual relative of *Liposcelis* sp., *Liposcelis bostrychophila*, are also maintained in the laboratory under the same conditions. We maintain three separate cultures of *L. bostrychophila* in the laboratory which we call 'AF', 'CGC' and 'MB'. The 'AF' culture was obtained from Kansas State University, 'CGC' was obtained from the Canadian Grain Council in Winnipeg Manitoba, and the 'MB' culture from Victoria BC.

Initial investigation into the putative horizontal gene transfer

I utilized the genome and transcriptome data available for nondistorter and distorter *Liposcelis* sp. (Perlman et al. 2015; Hamilton et al. 2018) to examine four genes that appeared to have a *Wolbachia* origin (Odile1-4). These genes had been previously identified in the distorter transcriptome and were identified as possible candidates for causing the sex ratio distortion in the distorters. I used Geneious 7.1.9 (Biomatters Ltd.) to set up Custom BLAST to the *Liposcelis* sp. genomes and transcriptomes (compiled data are available at the Dryad data repository) to find the genomic region of these genes and to look for related sequences. Additionally, I created alignments of the genomic regions and translated sequences using Geneious Alignment in order to examine the relatedness of the genes. I joined forward and reverse paired end RNA raw reads (100bp) from distorters in order to map them to reference sequences. Odile genes were used as reference sequences and sensitivity was customized. I disallowed gaps, set "maximum mismatches per read" to 0% and set the "maximum ambiguity" to 1 in order to more specifically map the raw reads to their respective gene copies. I wanted to confirm predicted coding sequence (CDS) of the putative horizontally transferred genes, as well as to look into the presence of introns.

I designed primers to amplify across the length of the horizontally transferred *Wolbachia* genes in the distorter genome in order to confirm the gene sequences with Sanger sequencing. I designed primers for a Nested PCR in order to improve the specificity of primers that bind to regions with very high similarity to the other genes

with a *Wolbachia* origin. I focus here on the amplification of one of the candidate genes, Odile1. The first primer set was designed to bind upstream and downstream of the CDS of Odile1 in a region of non-homology to the paralogs (Table S2-1). I used Q5 High-Fidelity DNA Polymerase (NEB) and followed the accompanying protocol to amplify the 4,148 bp product. I used the amplicons from the first PCR as a template for subsequent PCRs with the nested primer sets (Table S2-1). For these PCRs I used Taq DNA polymerase (ABM) and followed the accompanying protocol. Resulting amplicons were sequenced (Sequetech), and I aligned the overlapping regions of the amplicon sequences with each other to confirm the gene sequence (Geneious 7.1.9.). Additionally, I used a Custom BLAST in Geneious 7.1.9 to search for Odile homologs in the available transcriptome data from the parthenogenetic *Liposcelis bostrychophila* (Beibei strain) previously assembled from raw Illumina RNA read sets (PRJNA188391).

Confirmation of genomic integration

In order to confirm the putative HGT was indeed integrated into the distorter genome, and not a bacterial contamination, I carried out a number of experimental and bioinformatic analyses. Previous work on *Liposcelis* sp. did not reveal the presence of microbial symbionts; however, I re-screened both the *Liposcelis* sp. and *L. bostrychophila* to confirm our current lab cultures did not harbour *Wolbachia*. Screening was done with *Wolbachia* surface protein (wsp) specific primers (Table S2-2) and I used DNA from *Drosophila neotestacea* as a *Wolbachia* positive control. I mapped the transcripts of the putative HGTs to the associated genomic contiguous sequences (contigs) in the distorters to look for the possibility of introns, as well as to gather more information on the genome location where the putative integration occurred. Additionally, I compared predicted exons generated by the genome prediction pipeline to exons that I manually predicted in order to provide additional support for the presence and location of introns. I designed primers to bind to exon 1 and exon 2 of Odile4 in order to confirm the presence of the intron (Table S2-3). I used both genomic DNA and cDNA as a template for PCR, as well as two different extension times (10 sec and 45 sec) in order to compare the size of the products. I designed nested primers in order to amplify across the length of the intron in gDNA (Table S2-4). I visualized PCR products on a 1%

agarose gel (FroggaBio) following gel electrophoresis with the use of a 1kb plus DNA ladder (Invitrogen). I sent the cDNA amplicons to be sequenced (Sequetech) and analysed the resulting sequence in Geneious 7.1.9.

Phylogenetic analysis

To examine the phylogenetic relationship of the HGT event, I searched for homologs of the Odile gene in *Liposcelis* sp. distorters in non-redundant databases using BLASTp (Table S2-5). I included Odile homologs from the previously assembled transcriptome data for the asexual *L. bostrychophila* (Beibei), as well as amino acid sequence from the Alphaproteobacterium *Candidatus* Paracaedibacter acanthamoebae as the outgroup. I generated an alignment of amino acid sequences using Geneious alignment with automated parameterization in Geneious 7.1.9 (Biomatters Ltd.). I performed alignment character trimming with BMGE 1.12 (Criscuolo and Gribaldo, 2010) weighted with the BLOSUM35 similarity matrix and used Hyphy (Pond et al. 2005) to estimate the optimal model of amino acid substitution (WAG). I generated the phylogeny with PhyML 3.0 (Guindon et al. 2010) in SeaView 4.6.2, (Gouy et al. 2010), bootstrapped with 1000 replicates and visualized in Figtree 1.4.2 (Rambaut, 2007).

Investigation into the Odile homologs in Liposcelis bostrychophila

As Odile homologs were found in the transcriptome of *Liposcelis bostrychophila* (the asexual relative of *Liposcelis* sp.), I wanted to investigate whether these originated from a single horizontal gene transfer event, or whether they had been acquired independently. I designed degenerate primers based off of a DNA alignment of Odile genes and Odile homologs in the *L. bostrychophila* transcriptome (Table S2-6). I obtained amplicons of Odile homologs from two laboratory cultures of *L. bostrychophila*, 'AF' and 'CGC'. Additionally, I generated amplicons using the degenerate PCR primers from the *Liposcelis* sp. distorters and from a lab collection of *L. bostrychophila* (previously *L. granicola*) individuals that had been preserved in ethanol. I extracted DNA from pooled individuals by bead beating samples in lysis buffer (100mM Tris-HCL, 100mM NaCl, 50mM Na_2EDTA and 1%SDS) and a 1/10 volume of 3.3M NaOAc. I

incubated the samples for one hour at 65˚C followed by a phenol-chloroform extraction (chloroform was used at a 1/5 volume of the aqueous solution) and ethanol precipitation. Samples were sequenced (Sequetech) and additional sequences were obtained from the NCBI nucleotide collection (nr/nt) database (Table S2-7). I aligned the sequences in Geneious 7.1.9 using Geneious Alignment and performed alignment character trimming with BMGE 1.12 (Criscuolo and Gribaldo, 2010) with default parameters for DNA alignments. I used Hyphy (Pond et al. 2005) to estimate the optimal substitution model (model: 010110), and generated the phylogeny with PhyML 3.0 (Guindon et al. 2010) in Geneious 7.1.9, bootstrapped with 100 replicates and visualized in Figtree 1.4.2 (Rambaut, 2007).

Additionally, I looked at the evolutionary relationships of the booklice based on the cytochrome c oxidase I (COI) mitochondrial gene. I constructed a phylogeny in order to compare the relatedness of different *Liposcelis* species with which I could then look for evidence of either a single or an independent gene transfer event. COI sequences from *Liposcelis* species were obtained from the NCBI Nucleotide database (Table S2-8), as well as sequences generated from universal COI invertebrate DNA primers (Table S2-9). I amplified COI DNA from the *L. bostrychophila* (granicola) sample, the *L. bostrychophila* 'AF' and 'CGC' colonies, as well as a third lab colony of *L. bostrychophila* 'MB' from Victoria. Amplicons were sequenced (Sequetech) and I aligned the resulting sequences in Geneious 7.1.9 using ClustalW Alignment. The barklouse *Psococerastis albimaculata* was used as the outgroup. I performed alignment character trimming with BMGE 1.12 (Criscuolo and Gribaldo, 2010) with default parameters for DNA alignments and used Hyphy (Pond et al. 2005) to estimate the optimal substitution model (model: 012010). I generated the phylogeny with PhyML 3.0 (Guindon et al. 2010) in Geneious 7.1.9, bootstrapped with 100 replicates and visualized in Figtree 1.4.2 (Rambaut, 2007).

Conserved domain and putative function analysis

Homologs of Odile in the non-redundant (nr) database were annotated as uncharacterized proteins, providing little information on the role or function within its native *Wolbachia*. To try and gain more information on the potential function of the candidate HGT, I searched for conserved protein domains using the Conserved Domain Database (CDD) (Marchler-Bauer et al. 2017) on NCBI and InterPro: protein sequence analysis and classification (Hunter et al. 2009). I looked into other biological properties based on protein sequences in order to look further at the putative function of the candidate gene and its homologs. Using SignalP 4.1 (Peterson et al. 2011) and TMHMM 2.0 (Krogh et al. 2001), I looked for the presence of signal peptides and transmembrane helices. Additionally, I looked at the predicted subcellular protein localization using DeepLoc-1.0 (Armenteros et al. 2017), PSORTb (Yu et al. 2010), ESLpred (Bhasin and Raghave, 2010), BUSCA (Savojardo et al. 2018), CELLO (Yu et al. 2006) and SOSUI-GramN (Imai et al. 2008).

Life stage expression data

In order to supplement existing RNA seq data, I wanted to look at transcript abundance of Odile1 using RT-PCR. Specifically, I wanted to see if there was a difference in transcript abundance at different life stages of the *Liposcelis* sp. distorters. I extracted total RNA from individual distorter *Liposcelis* sp. at four different life stages, 3 day old nymphs (n = 6), 12 day old nymphs (n = 6), 49 day old un-mated adults (n = 4) and 52 day old mated adults (n = 6) as well as nymphs that were already morphologically female but not yet adults (approximately 1 month old) (n = 4). To obtain mated adults, I paired distorters with males in individual dishes and waited until there was the presence of nymphs in each dish (14 days). Total RNA was isolated from all individuals with 50 uL of TRIzol (Invitrogen) and 1/5 volume of the aqueous phase of chloroform. I also added 0.01 mg RNA grade glycogen (Thermo Scientific) to aid with RNA ethanol precipitation and pellet visualization. RNA was precipitated at -80˚C for 45 min in 95% ethanol, washed with 80% ethanol and stored in 80% ethanol overnight at -20˚C. I removed the final ethanol wash, dried the pellet via evaporation and re-dissolved the RNA in 11uL UltraPure Distilled Water (Invitrogen). I quantified and assessed the purity

of samples by Nanodrop absorbance readings. All centrifugation steps were carried out at 4˚C at 14k rpm (Hettich Instruments, Mikro 120).

I removed genomic DNA from RNA for each sample using RNase free DNase I (ThermoScientific). Samples were DNase I treated for 30 min at 37˚C and the reaction was terminated with the addition of 50mM EDTA at 65˚C for 10 min. To obtain cDNA I used the SuperScript II Reverse Transcriptase kit (Invitrogen) according to the manufacturer's protocol. I primed each sample for reverse transcription with 1 uL of anchored Oligo$(dT)_{20}$ primer (50uM) (IDT). Additionally, I included "no RT" controls for each sample that did not undergo reverse transcription in order to determine the presence of genomic DNA contamination.

I performed real-time qPCR with a CFX96 Real-Time System (BioRad) with SYBR Select Master Mix (Applied Biosystems). Each reaction contained 2 µL of 1:10 diluted cDNA template, 5 µL of SYBR Select master mix 1.4 µL UltraPure distilled water and 0.8 µL of each primer (5 µM) (Table S2-10). Each sample was run in triplicate with Odile1 primers and in duplicate for GAPDH and RPL0 primers. All samples were run on 96-well plates and I included interplate calibrators on each plate to control for variation between qPCR runs. I calculated the relative expression of Odile1 at each life stage with a modified Pfaffl equation (Pfaffl, 2001), see Equation 1.

$$[(E_{target})^{\Delta Ct(\text{gm Odile1 - sample})}] \,/\, [(E_{ref})^{\Delta Ct(\text{gm ref - sample})}] \quad (1)$$

I calculated the global mean (gm) of the C_t values for each gene target to use as controls as I was not looking at treated versus untreated samples. I obtained real-time PCR efficiencies (E) from the given slopes of dilution series in BioRad CFX Manager 3.0. I averaged technical replicates that were within 0.5 C_t of each other, and I used the average C_t of both reference genes for each sample.

I analyzed all data in R Studio v1.0.136 (R Core Team, 2016). I used a non-parametric analysis after considering the sample size and the non-normality of the data determined with a Shapiro-Wilk test. I used a Kruskal-Wallis test by ranks to see if there

was a significant difference in relative transcript abundance across life stages. In order to look at post-hoc pair-wise relationships, I used a Kruskal-Nemenyi test with Tukey distribution for independent samples.

Results

Initial investigation into the putative Wolbachia HGT event

In addition to the three previously identified genes of *Wolbachia* origin in *Liposcelis* sp. (Odile1, Odile2 and Odile3), I found a fourth gene (Odile4) using a custom BLAST of the distorter genome. This gene was not originally annotated as having a *Wolbachia* origin as the first BLAST hit in the nr database is from *Wasmannia auropunctata*, the little fire ant, and not *Wolbachia*; however, the subsequent hits show a *Wolbachia* origin. I mapped the genes to their locations on the genomic contigs and found that three of the genes, Odile1, Odile3 and Odile4, mapped to one contig 77 kb in length while the fourth, Odile2, mapped to a second contig 91 kb in length. I aligned the four genomic sequences (including intronic regions) of *Wolbachia* origin and saw that they were in fact paralogs and align with a 93.8% pairwise identity. An amino acid alignment of the four translated CDS had a pairwise identity of 75.7% (Figure 2-1).

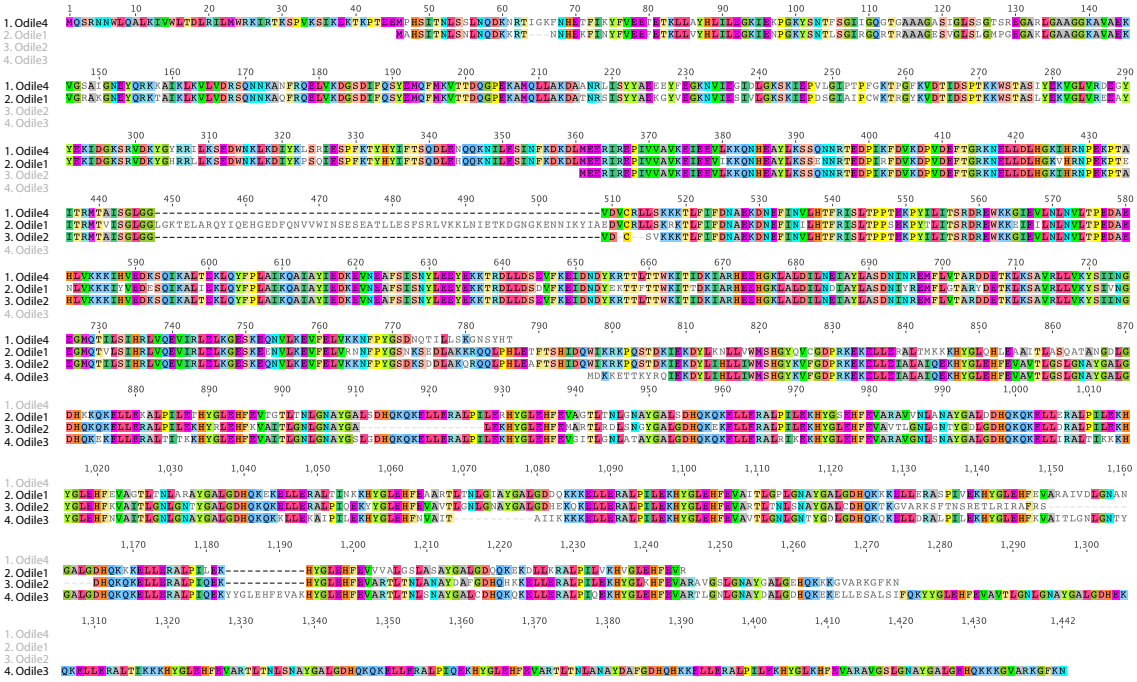


Figure 2-1. Amino acid alignment of the four paralogs in the distorter *Liposcelis* sp. genome of *Wolbachia* origin. Alignment was performed using Geneious Alignment (BLOSUM62) in Geneious 7.1.9 and has an average pairwise identity of 75.7%. Gaps in the amino acid sequences are indicated with a dash (-) and highlighted residues represent a match to the consensus sequence. Amino acid positions within the alignment are numbered.

The previously predicted CDS for Odile1 from the distorter transcriptome appeared to be truncated based on open reading frame predictions and translation in Geneious 7.1.9. The original CDS was predicted to be 3,229 bp and contained putative introns in the 5' untranslated region. Raw RNA reads mapped to the genomic sequence also supported a longer CDS (3,561 bp) and I continued my analysis of Odile1 with the assumption that this was the full length of the gene. Odile1 has the most complete coding region out of the paralogs and for this reason I focused primarily on this gene as the candidate for sex ratio distortion. As well, it appears as if Odile4 and Odile3 were originally one gene, as they are located 27 bp apart in the assembled genome and each align to Odile1 with no overlap between them. Now, each gene has a unique open reading frame and different transcript levels (based on the previously generated RNAseq data).

Using PCR, I amplified the full length of Odile1, followed by nested PCR with primers that spanned across the gene. I sequenced the resulting amplicons and was able to

confirm 3,537 bps with Sanger sequencing (MH751905), but was unable to generate a sequence for the last 24bps. I found that the Odile genes had been integrated into a region that is highly repetitive across the genome. This, paired with the copies of the Odile gene and the repeat motifs present in the Odile genes, created challenges when trying to design specific primers or isolate specific regions of the genes.

Custom BLAST searches in the *L. bostrychophila* transcriptome revealed eleven homologs of Odile, nine of them ranging from 315 bps to 809 bps, while the other two were 2,962 bps and 2,893 bps in length. Aligned with Odile, the two homologs had an average pairwise identity of 67.4% and 67.8% respectively, and had an average pairwise identity to each other of 86.0%.

Confirmation of genomic integration

In order to confirm the putative *Wolbachia* genes were not a result of endosymbiotic contamination, I screened lab samples of booklice with *Wolbachia* specific primers and did not find any evidence of *Wolbachia* endosymbionts in the lab cultures. After aligning the Odile transcripts to genomic contigs, I discovered the presence of spliceosomal introns. I looked for signature splice sites at each intron-exon junction and found that each contained the signature 5' splice donor site of GT and 3' splice acceptor site of AG. Considering bacteria lack spliceosomal introns, this provides further evidence of gene integration rather than bacterial contamination in the distorter *Liposcelis* sp. genome. Predicted introns are depicted in Figure 2-2. with intron length in nucleotides noted. Odile1 does not contain any introns in the coding region, but there are predicted introns in the 5' untranslated region. Odile2 contains four introns in the coding region, Odile3 contains two predicted introns in the coding region, and Odile4 has one predicted intron in the coding region and one in the 5' UTR (Figure 2-2B).

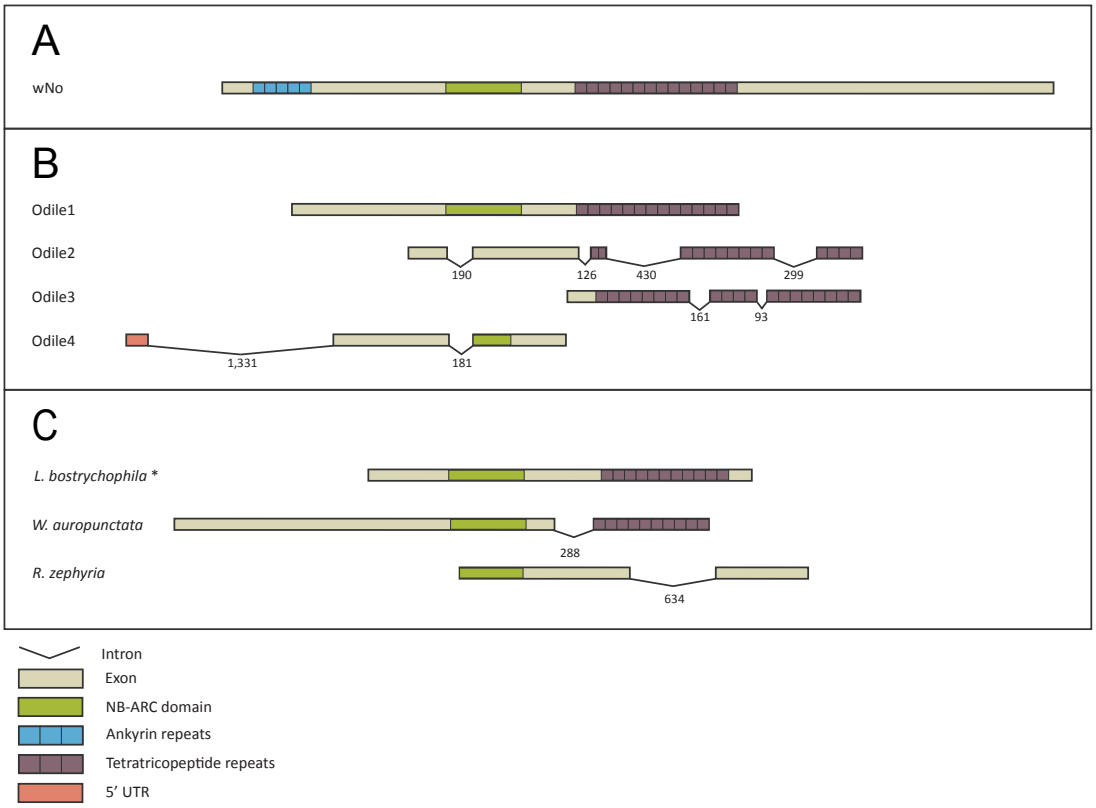


Figure 2-2. Nucleotide alignment schematic of the candidate gene homologs and paralogs showing introns and conserved domain locations. **A.** *Wolbachia* homolog from *Drosophila simulans* (wNo). **B.** Candidate gene (Odile1) and paralogs (Odile2, Odile3 and Odile4) in *Liposcelis* sp. distorter. **C.** Homologs found in the insects *Liposcelis bostrychophila*, *Wasmannia auropunctata* and *Rhagoletis zephyria*. Relative domain positions were predicted by CDD. Numbers below introns represent length in nucleotides. *Intron presence/absence is unknown for *L. bostrychophila* as only transcript sequence was available.

The predicted exons from the genomic pipeline and the exons I manually recovered by aligning transcripts to genomic regions were the same for Odile2, Odile3 and Odile4 providing additional evidence for the presence of introns in these genes. I was able to confirm the presence of the first intron in Odile4 via PCR and Sanger sequencing. The genomic DNA amplicon visualized on a gel appeared to be around 1,500 bp, while the cDNA amplicon appeared to be around 350 bp. The size difference between the two amplicons represents an un-spliced (gDNA) sample and a spliced (cDNA) sample (Figure 2-3). Sanger sequencing confirmed the genomic sequence from the genome assembly and the sequenced cDNA product confirmed the presence of the intron. However, the intron

size was different than what was predicted by the transcriptome and genome prediction pipeline. Sanger sequencing confirmed that the intron size was in fact 1,331 bp rather than the predicted 1,465 bp. As well, the sequencing results changed the predicted open reading frame of Odile4 excluding exon1, which is now predicted to be untranslated. The putative intronic regions of the Odile genes are of *Wolbachia* origin and align to exonic regions of the other paralogs. This suggests that the introns were not introduced, rather the spliceosomal machinery of *Liposcelis* sp. is recognizing erroneous splice sites in the bacterial genes and splicing out the DNA sequence between them.

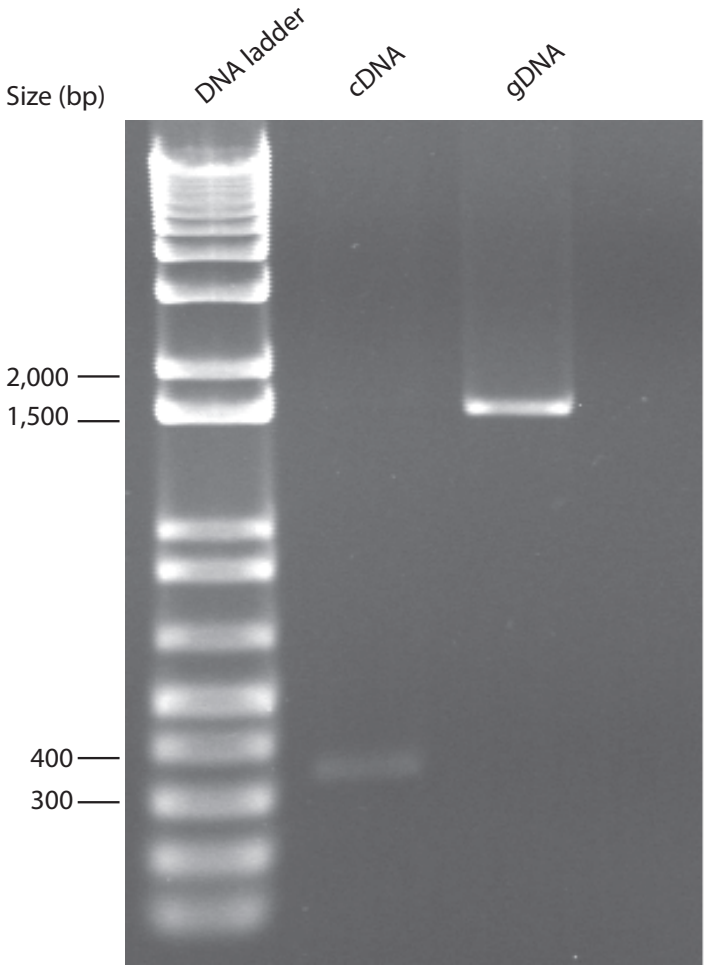


Figure 2-3. RT-PCR product (cDNA) and genomic DNA (gDNA) product from Odile4. Primers used were designed to bind to exon 1 and exon 2 in order to amplify across the putative intron. The bands correspond to the spliced cDNA product and the unspliced gDNA product. PCR products and a 1kb plus DNA ladder were separated on a 1% agarose gel.

Providing further evidence of genomic integration, the candidate gene and paralogs are located on two large genomic contigs, three of which (Odile1, Odile3 and Odile4) are located on a contig (Backbone_1352) 77,756 bp in length and one of which (Odile2) is located on a contig (Backbone_1144) 91,660 bp in length. Neither genomic contig contained any other bacterial genes, and a BLASTn search revealed hits to three insect genes on the genomic contig 91k bp in length.

Phylogenetic analysis

I found that the candidate gene for the sex ratio distortion in *Liposcelis* sp. was most closely related to predicted *Wolbachia* proteins (Figure 2-4). Similarly, the homolog from the closely related asexual *L. bostrychophila* is found within the same clade. Interestingly, I found two additional insect encoded proteins that also group within this *Wolbachia* clade. I did not consider the insect genes from *Wasmannia auropunctata* and *Rhagoletis zephyria* to be bacterial contamination as I found an intron in each gene (Figure 2-2C) and both assembled on large genomic contigs (128 kb and 5.9 Mb).

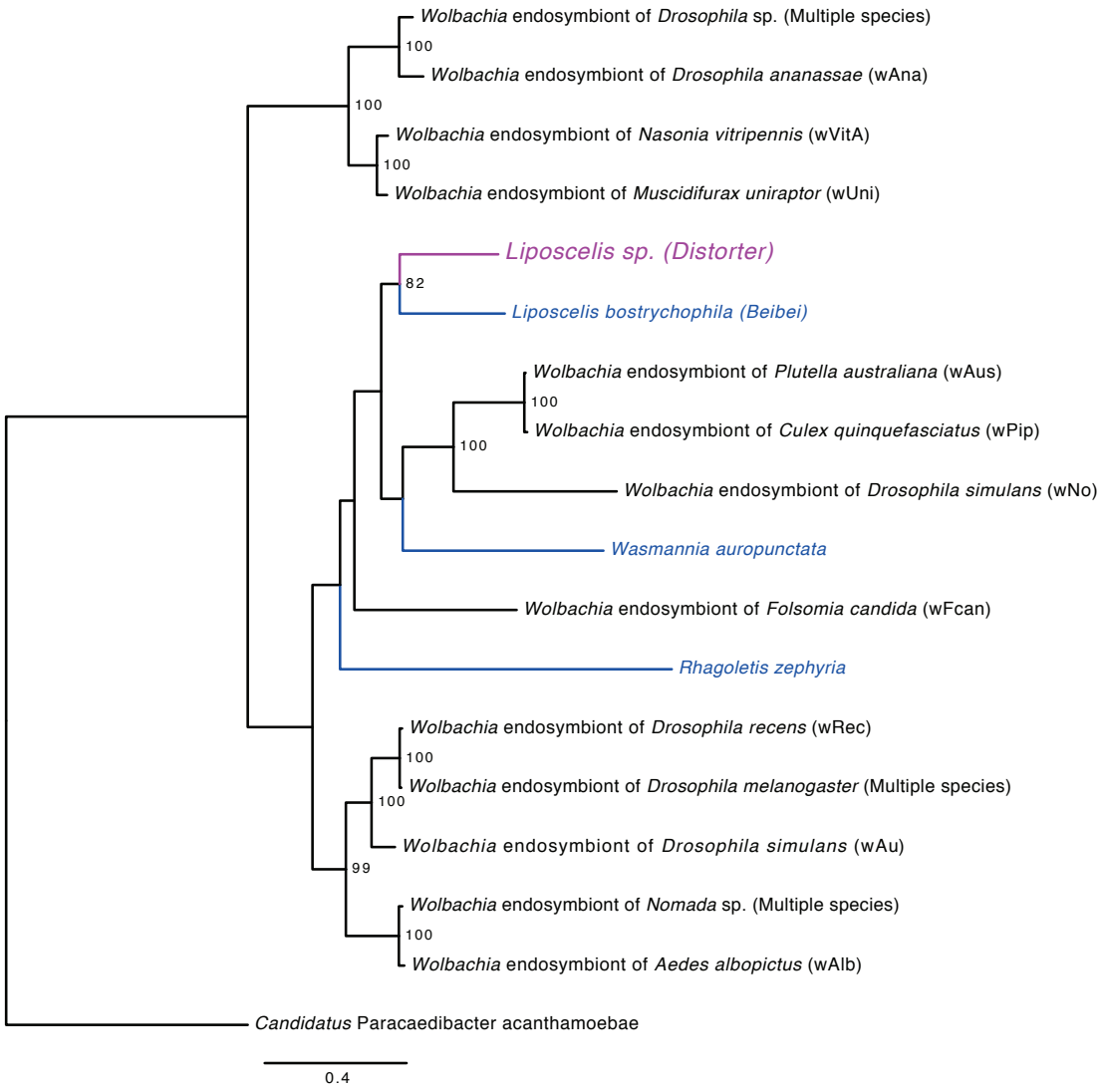


Figure 2-4. Maximum likelihood phylogenetic tree based on amino acid sequence of Odile1 in *Liposcelis* sp. distorters. Odile1 sequence is highlighted in purple, and sequences from putative HGT events to insect genomes are highlighted in blue. Odile1 homolog sequences came from the non-redundant database in NCBI. Nodal support was generated with 1,000 bootstrap replicates. Support less than 75 is not shown. Scale bar represents substitutions per site.

Investigation into the Odile homologs in Liposcelis bostrychophila

The *Liposcelis* species form a single clade in the maximum likelihood phylogeny of the Odile homolog nucleotide sequences (Figure 2-5). From this phylogeny alone, it is impossible to determine whether the *Wolbachia* gene had been acquired independently in each species or not.

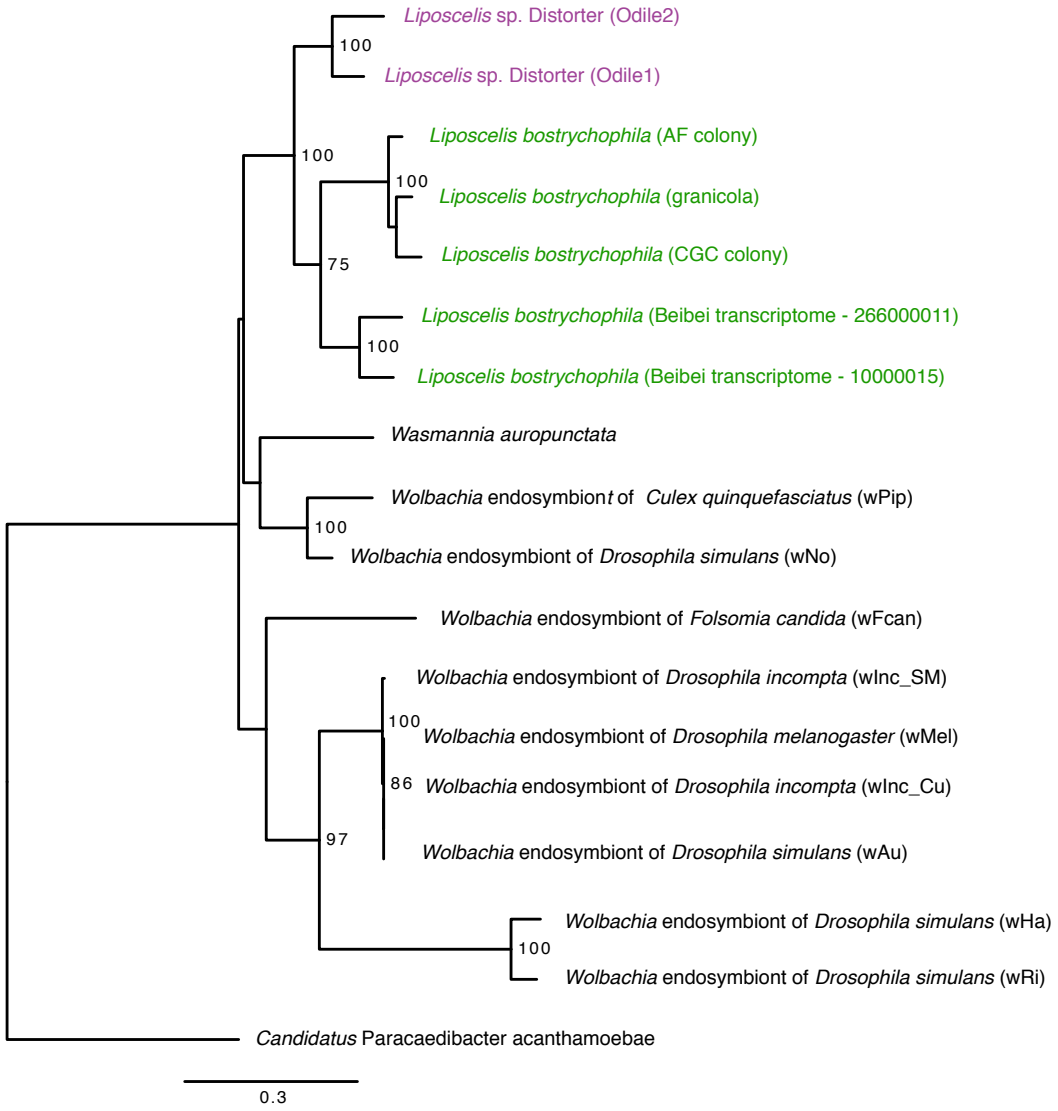


Figure 2-5. Maximum likelihood nucleotide phylogeny of the Odile genes (purple) including homologs from the asexual relative *Liposcelis bostrychophila* (green). Sequences from *L. bostrychophila* 'AF', *L. bostrychophila* 'CGC' and *L. granicola* result from Sanger sequencing, *L. bostrychophila* 'Beibei' sequences are homologs from the assembled transcriptome. Nodal support was generated with 100 bootstrap replicates. Support less than 75 is not shown. Scale bar represents substitutions per site.

In order to resolve the history of the gene transfer event, I generated a phylogeny of the evolutionary relationship of the *Liposcelis* species based on COI mitochondrial gene sequences. The *L. bostrychophila* species cluster together, and although there is low bootstrap support for the node, this clade also contains three sexual booklice species, *L.*

corrodens, *L. sculptilimacula* and *L. paeta*. Currently there is not sequence data available from these three species to determine if there are Odile homologs present; however, a PCR screen with degenerate primers (Table S2-5) of ethanol-preserved *L. corrodens* and *L. paeta* individuals did not result in DNA amplification. In this case, it is more parsimonious to assume that the *Wolbachia* gene was independently transferred two times, rather than transferred once and lost three times.

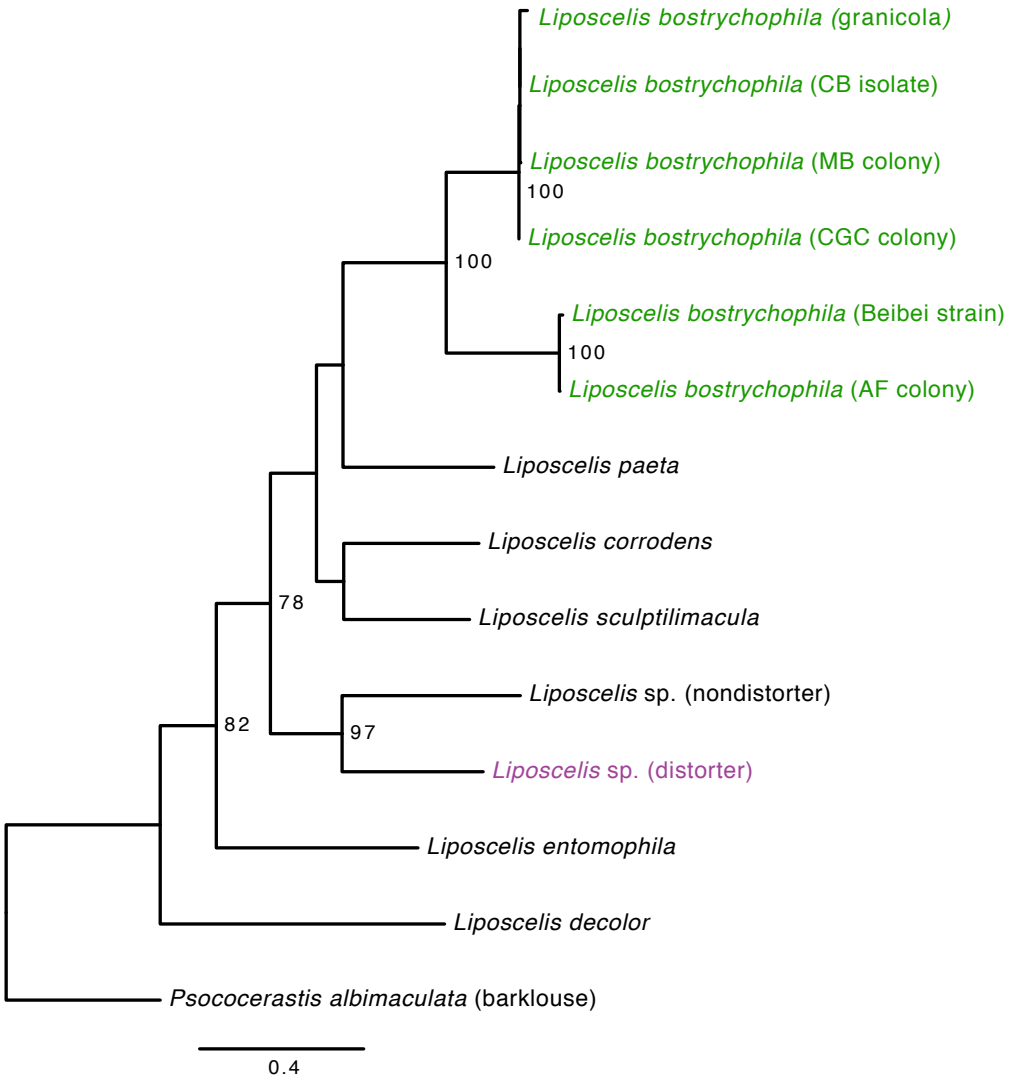


Figure 2-6. Maximum likelihood nucleotide phylogeny constructed from cytochrome c oxidase subunit I (COI) mitochondrial gene sequences. Sequences from *L. bostrychophila* 'AF', *L. bostrychophila* 'CGC' and *L. bostrychophila* 'granicola' result from Sanger sequencing. Nodal support was generated with 100 bootstrap replicates. Support less than 75 is not shown. Scale bar represents substitutions per site.

Conserved domain and putative function analysis

The *Wolbachia* homologs of the distorter gene candidate codes for ankyrin repeat motifs at the 5' end, as well as a NB-ARC domain, and tetratricopeptide repeat (TPR) motifs (Figure 2-4A). Similarly, I found that the Odile1 gene in the distorter genome contained an NB-ARC domain and TPRs at the 3' end, but was lacking the ankyrin repeat domain at the 5' end. The additional gene copies in the distorter genome contained combinations of the conserved domains. Odile2 contains TPRs but has acquired an intron that interrupts the coding region for the NB-ARC domain, Odile4 contains a partial NB-ARC domain and Odile3 contains TPRs (Figure 2-4B). I also examined the conserved domains in the homologs of the three other insect genomes. *Liposcelis bostrychophila* and *Wasmannia auropunctata* contained both the TPR domains and NB-ARC domain, while *Rhagoletis zephyria* contained only the NB-ARC domain (Figure 2-4C). Ankyrin repeat domains are only present in the *Wolbachia* homologs and were not found in any of the insect genes.

Neither the distorter candidate gene nor its homologs were predicted to contain a signal peptide or transmembrane helices by SignalP and TMHMM. Additionally, the distorter protein was predicted by DeepLoc-1.0, ESLpred and CELLO to have a cytosolic subcellular localization, while the *Wolbachia* (wNO) homolog was predicted by PSORTb, SOSUI-GramN, BUSCA and CELLO to have an outer membrane localization.

Reference gene selection for relative quantification in real-time qPCR

I tested primers for multiple reference genes in order to ensure stable expression across life stages. I checked that each reference gene was present as a single copy by using a custom BLAST of the distorter transcriptome, and designed primers based on distorter transcript sequence for GAPDH, RPL32, RPL0, Trf2, Ef1a (Table S2-10) (IDT). I performed an initial trial of the reference gene primers using real-time qPCR with a CFX96 Real-Time System (BioRad) in triplicate across three life stages, nymph, pre-adult and adult. I found no significant differences in threshold cycle (C_t) across life stages for any of the reference genes (Kruskal-Wallis chi-squared = 2, df = 2, p-value = 0.37). I chose the two reference genes with the lowest standard deviations (GAPDH and RPL0)

to use for normalization of the target gene. Following qPCR of the two reference genes and of Odile1 across the five life stages, I compared the reference primers across life stages in order to confirm there were no significant differences, and to remove outliers. Outliers seen in Figure 2-7 for reference genes GAPDH (Figure 2-7A) and RPL0 (Figure 2-7B) were omitted from the relative quantification calculations.

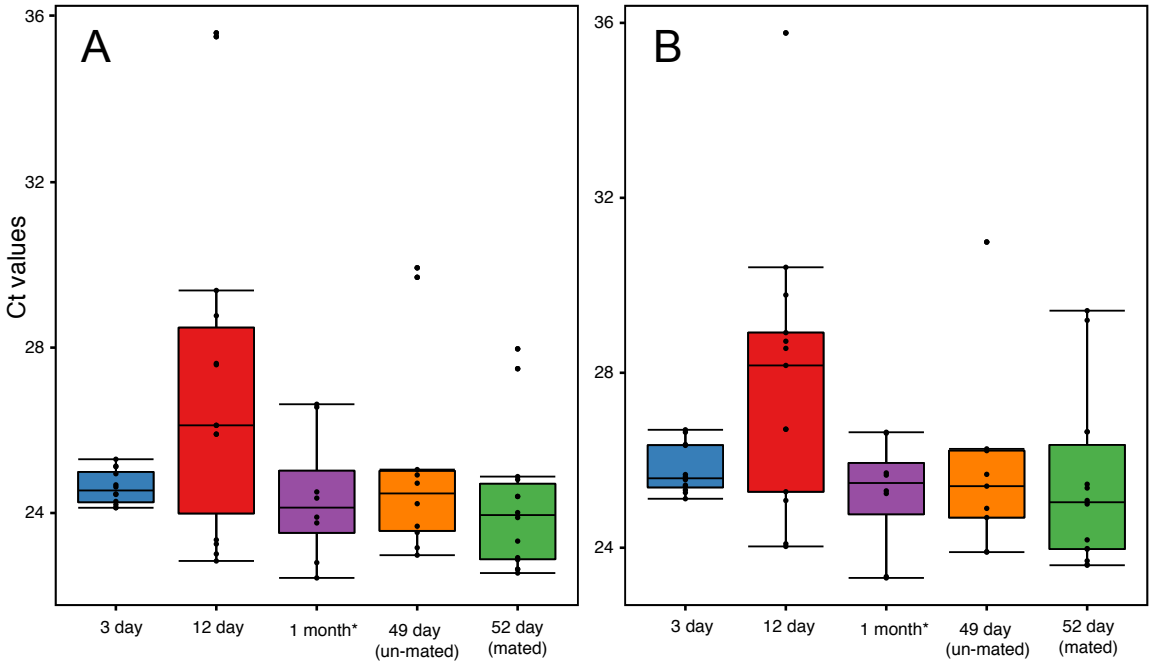


Figure 2-7. Analysis of reference genes used to normalize Odile1 relative expression. **A.** Boxplot of GAPDH cycle threshold (Ct) values across five life stages of *Liposcelis* sp. distorter from RT-qPCR analysis. **B.** Boxplot of RPL0 Ct values across five life stages of Liposcelis sp. distorter from RT-qPCR analysis. Life stages were from 3 day-old nymphs (n = 12), 12 day-old nymphs (n = 14), nymphs that were morphologically female (approximately 1 month) (n = 8), 49-day old unmated adults (n = 10) and 52 day mated adults (n = 14). Outliers are denoted as points above the upper quartile. * Individuals that are approximately 1 month old are not yet adults but are morphologically female.

With outliers omitted, I fit an Analysis of Variance (ANOVA) model for each reference gene to check that there was no significant difference in Ct across the five life stages. Neither GAPDH ($F(4,30) = 1.341$, $p = .278$) nor RPL0 ($F(4,29) = 1.765$, $p = 0.163$) showed a significant effect of life stage.

Life-stage specific transcript abundance of Odile1

As a way to further investigate putative function, as well as provide more evidence toward nuclear genome integration, I looked into life-stage specific gene expression of Odile1. Relative transcript abundance of Odile1 to the global mean was calculated across five life stages, 3 day-old nymphs, 12 day-old nymphs, nymphs that were morphologically female (approximately 1 month), 49-day old unmated adults and 52 day mated adults. All gene expression values were normalized using two reference genes and Log2 transformed (Figure 2-8).

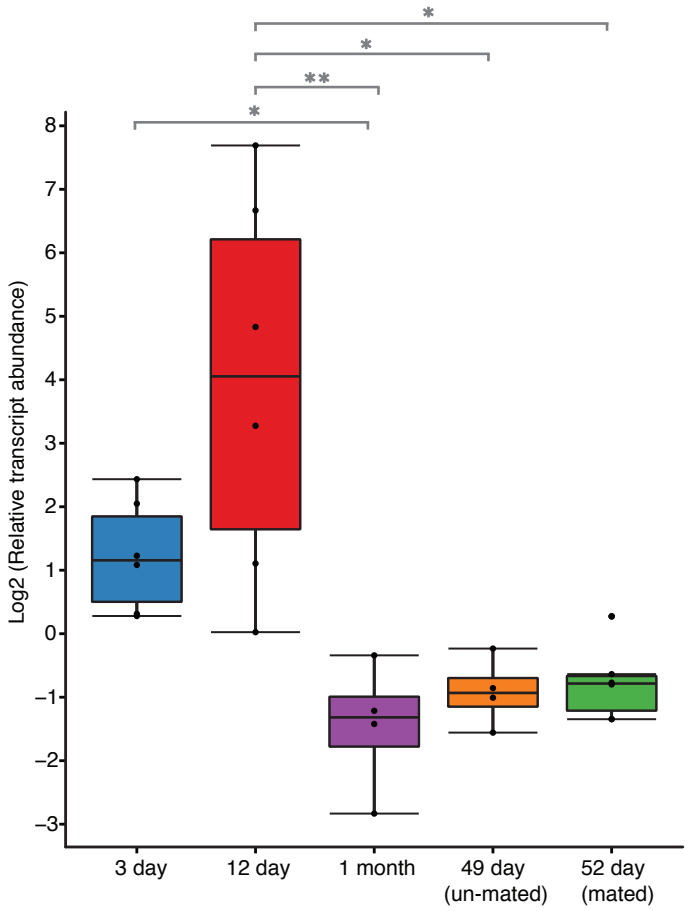


Figure 2-8. Log2 relative gene expression of Odile1 across five life stages of distorter *Liposcelis* sp., 3 day-old nymphs (n = 6), 12 day-old nymphs (n = 6), nymphs that were morphologically female (approximately 1 month) (n = 4), 49-day old unmated adults (n = 4) and 52 day mated adults (n = 6). Expression values are reported relative to the global mean of Odile1 and normalized with GAPDH and RPL0 reference genes. * Individuals that are approximately 1 month old are not yet adults but are morphologically female.

In general, transcript abundance for the Odile1 was low across all life stages, with a global mean of 33.69 C_t, compared to the reference genes that had a combined global mean of 24.82 C_t. Even at relatively low transcript abundance, there appears to be a higher relative abundance of Odile transcripts in the 3 day-old samples and the 12 day-old samples compared to the other life stages. The Kruskal-Wallis test showed a significant difference present across the life stages (χ^2 (4, N = 30) = 6.92, p = < .001), but does not specify which groups the significant difference occurs between. I conducted a post-hoc analysis with the Kruskal-Nemenyi test to examine pair-wise relationships between the different life stages (Table 2-1). Significant differences were found between the 12 day-old group and the 1-month group ($p < .01$), and between the 12 day-old group and the un-mated and mated groups ($p < .05$). As well, there was a significant difference between the 3 day-old group and the 1-month group ($p < .05$).

Table 2-1. Kruskal-Nemenyi pair-wise comparisons and statistics of relative Odile1 transcript data across life stages. Relative gene expression was determined through qPCR on cDNA of age-matched distorter *Liposcelis* sp. individuals, normalized with GAPDH and RPL0 reference genes relative to the global mean of Odile1 cycle thresholds. Groups include distorter individuals at 3 days-old, 12 days-old, approximately 1 month old, 49 days-old (un-mated) and 52 days-old (mated).

					Kruskal-Nemenyi pair-wise comparisons			
Group	*n*	Mean	Median	SE	3 day	12 day	1 month	Un-mated
3 day	6	2.74	2.23	0.687	-	-	-	-
12 day	6	58.3	19.1	33.4	.097	-	-	-
1 month	4	0.434	0.402	0.134	.049*	.0083**	-	-
Un-mated	4	0.560	0.524	0.107	.14	.030*	1.0	-
Mated	6	0.633	0.580	0.123	.16	.030*	.95	1.0

* Significant at $p < 0.05$, ** significant at $p < 0.01$

Discussion

There are many examples of organisms with sex ratios that deviate from an equal ratio of male to female offspring. These deviations can be caused by selfish genetic elements that have evolved the ability to manipulate reproduction in order to increase their transmission to the next generation. In *Liposcelis* sp. a selfish genetic element has caused an extreme sex ratio distortion in which one female type (distorter females) will produce exclusively female offspring. A candidate gene, Odile, has been identified in the genome of the distorter females and is proposed to be the selfish genetic element responsible for the observed sex ratio distortion. Here, I investigate and characterize Odile through both bioinformatics and experimental techniques. I present evidence that the candidate gene of *Wolbachia* origin is integrated into the genome of the distorter females, and that it has been duplicated a number of times. I report the presence of conserved domains in the gene of interest, along with other biological properties. Additionally, I examine the evolutionary history of Odile through phylogenetic analysis, and the expression pattern across different life stages of distorter females.

Horizontal gene transfers from prokaryotes to eukaryotes are becoming more widely reported with most cases thus far involving *Wolbachia* endosymbionts (Werren et al. 2008). The *Wolbachia* gene found in the genome of the distorter *Liposcelis* sp. booklouse represents a candidate for the sex ratio distortion found in this species. It is likely that this gene stemmed from an initial lateral transfer event followed by duplication. Although the genomic sequence similarity is high, with a pairwise identity of 93.8%, the four Odile genes have different coding sequences (CDS). Similarly, in the parasitoid wasp *Nasonia vitripennis*, thirteen genes of *Wolbachia* origin were identified in the nuclear genome (The Nasonia Working Group, 2010). It was predicted that one or more of the genes had been horizontally transferred from *Wolbachia* and subsequently amplified and diverged in the genome of the wasp.

In order to preclude bacterial contamination as a source of the candidate genes, I screened current lab cultures for *Wolbachia* endosymbionts and did not see any evidence *Wolbachia* in the cultures. This, taken with previous work to screen this species with targeted and untargeted molecular screens, as well as microscopy, provides convincing evidence that these genes do not represent bacterial contamination. Additionally, the

genes are located on large contigs (77 kbp and 91 kbp in length) that do not contain any other bacterial genes. The fact that there are no other bacterial genes on these two large genomic contiguous sequences is significant as most prokaryotic genomes consist of more than 85% coding DNA (Lynch, 2006); therefore, if the contiguous sequence was bacterial contamination, we would not expect to see long stretches of non-coding DNA as we see in these large genomic contigs.

The presence of introns provides another piece of evidence that supports the integration of the *Wolbachia* genes into the genome of *Liposcelis* sp. distorters. Transcript alignments to genomic sequence revealed the presence of putative introns with signature splice sites. Because spliceosomal introns are not found in bacteria, the presence of these type of introns indicates that the genes are eukaryotic (Rogozin et al. 2003). The presence of spliceosomal introns have provided evidence toward horizontal gene transfer from bacteria to eukaryotes previously in bdelloid rotifers, small, freshwater invertebrates, where at least two genes showing a bacterial origin had acquired an intron (Gladyshev et al. 2008). Similarly, spliceosomal introns were identified in the horizontally transferred bacterial genes in the plant-parasitic nematode genome (Craig et al. 2008; Danchin et al. 2010).

Phylogenetic analysis supports a *Wolbachia* origin for the Odile candidate gene found in the *Liposcelis* sp. distorter genome. The distorter gene maps within a clade of *Wolbachia* genes, along with the homolog found in the related booklouse *Liposcelis bostrychophila*. Two additional insect genes grouped within the *Wolbachia* clade, *Wasmannia auropunctata* and *Rhagoletis zephyria,* suggesting that the *Wolbachia* gene has been independently integrated into at least three insect genomes. Interestingly, *L. bostrychophila* and *W. auropunctata* also have unusual modes of reproduction. In *L. bostrychophila*, a common stored grain pest, reproduction occurs through parthenogenesis in which embryos develop without fertilization (Shires, 1982), while in the little fire ant *W. auropunctata*, both males and queens are produced clonally (Fournier et al. 2005). There is no gene flow occurring between the male and queen gene pools as worker ants produced through normal sexual reproduction are sterile.

A phylogeny showing the evolutionary history of *Liposcelis* species based on the nucleotide sequence of the mitochondrial gene COI, provides some additional evidence that the *Wolbachia* genes found in both *Liposcelis* species originate from independent transfer events. *Liposcelis bostrychophila* and *Liposcelis* sp. are not closely related (Feng et al. 2018; Figure 2C). Odile genes were not detected in the transcriptome of *L. entomophila* (Hamilton et al. 2018), and PCR screens with degenerate primers for Odile homologs in *L. paeta*, *L. corrodens*, *L. decolor, L. entomophila* and nondistorter *Liposcelis* sp. did not result in amplification. Based on this preliminary search for Odile homologs in related booklouse species, multiple gene loss events would have had to take place if the presence of *Wolbachia* genes in these two booklouse originated from a single gene transfer event. However, further sampling and sequencing still needs to be done in order to resolve these phylogenetic trees and provide further bootstrap nodal support.

The bacterial homologs of the Odile genes were uncharacterized in the NCBI RefSeq database, but were often annotated according to their conserved domains, such as ankyrin repeat domains or tetratricopeptide repeat (TPR) motif. The Odile genes in distorters did not contain ankyrin repeat domains found in their *Wolbachia* homologs, but did contain TPRs and an NB-ARC domain. Neither domain provides a solid indication of putative function as they are documented to be involved in various biological processes. The TPR motifs are a degenerate, 34 amino acid sequence present in tandem arrays of 3-16 motifs that form anti-parallel α-helices and are involved in mediating protein-protein interactions (Das et al. 1998). Present in a wide range of functionally diverse proteins, TPRs are often associated with multiprotein complexes, and have been found to be important in the functioning of molecular chaperone complexes, anaphase promoting complexes, transcription repression complexes and protein import complexes (Blatch and Lässle, 1999). The ankyrin repeat motif present in the *Wolbachia* homologs but absent from the Odile genes are also involved in protein-protein interactions and have a similar structure to TPRs but also contain β-strands (Groves and Barford, 1999). The NB-ARC domain is known to hydrolyse ATP and be involved in programmed cell death as a component of resistance (R) genes in plant defense (van Ooijen et al. 2008). The plant resistance genes share homology with cell death genes in *Caenorhabditis elegans* (CED-4) and a human apoptotic protease-activating factor (APAF-1) (van der Biezen and Jones,

1998). Interestingly, these cell-death related proteins containing a NB-ARC domain are all followed by repeat regions involved in protein-protein interactions similar to TPRs. The Odile homologs found in *L. bostrychophila* and *W. auropunctata* also contain the NB-ARC domain and tetratricopeptide repeat motif. The snowberry fruit fly, *R. zephyria* that is not known to have an unusual form of reproduction, does not contain TPRs as a putative intron is located in this region.

Looking into other biological characteristics of a gene can provide additional insight into putative protein function, such as predicted localization of a protein or whether there are protein sorting signals present or not. Signal peptides, located at the N-terminus of the amino acid chain, prompt the translocation of a protein to specific membranous regions, such as the outer membrane or certain organelles (Izard and Kendall, 1994). Neither the Odile1 gene nor the *Wolbachia* gene homologs contained signal peptides, suggesting a cellular localization within the cytosol absent any further targeting sequences. Further localization analysis of Odile1 predicted that it is a cytosolic protein, but is not more specific than that. Interestingly, the homolog in the *Wolbachia* endosymbiont of *Drosophila simulans (*wNo*)* and in *Culex quinquefasciatus* (wPip) were predicted to be outer membrane proteins, despite the absence of a predicted signal peptide.

The Odile genes were previously shown via RNAseq to be transcribed, although at relatively low levels. Even though transcription occurs, this does not imply functionality (Dunning Hotopp, 2011). Determining functionality of HGTs can be difficult without an experimentally determined phenotype, but previous studies have shown that tissue or life stage-specific transcription may be an indication. In two distantly related nematode species (*Acanthocheilonema viteae* and *Onchocerca flexuosa*), tissue-specific expression of *Wolbachia* genes was found in developing embryos and testes (McNulty et al. 2010). As well, in another nematode species (*Brugia malayi*), three *Wolbachia* genes are differentially transcribed in a life-stage specific manner (Ioannidis et al. 2013). As constitutive expression of horizontally transferred genes at low levels is not uncommon, tissue or life-stage specific transcription can be an indicator of potential functionality (Dunning Hotopp, 2011).

Following RT-qPCR, the Odile1 gene exhibited low levels of transcription at all life stages; however, there were significant differences in transcript abundance between early life stages (3 days and 12 days) and older life stages (>1 month). Before reaching adulthood, *Liposcelis* sp. nymphs, whether male or female, are indistinguishable. After approximately one month, females are distinguishable from males by being larger in size with a larger abdomen. It is interesting to see that a more variable transcript level of Odile1 occurs before the distorters are morphologically female, suggesting Odile1 could play a role in development or sex determination. In *Liposcelis* sp., sex determination is through paternal genome elimination in which paternal chromosomes in males are silenced via heterochromatinization (Hodson et al. 2017). The sex of the offspring is therefore determined by whether or not the paternal chromosomes are inactivated in the egg following fertilization (Haig, 1993). In *Liposcelis* sp., the distorter females are acting similarly to males, in which only the maternal genes are passed on to the next generation. In order to better speculate on the mechanism of distortion, it would be beneficial to know at what developmental stage the paternal genome is being eliminated, whether it occurs following fertilization, or early in development. The underlying mechanism involved in the elimination of the paternal genes from the distorter germline is unknown, but as the transcript level of Odile1 is increased in early development, it would be interesting to see if it has a role in this process. Additionally, it is unknown at what transcript abundance Odile1 gene expression would be biologically relevant, but the age-specific transcript abundance could provide further evidence toward putative function.

Chapter 2 – Supplemental Information

Supplemental Table 2-1. Primers and thermocycling conditions used to amplify the full length of the Odile1 gene and the subsequent nested primers and thermocycling conditions used to amplify across the Odile1 gene.

Primer name	Sequence	PCR conditions
6313p2F	GATCATATGCCACTGTTTACTGC	98°C × 30sec, (98°C × 10sec, 63°C × 30sec, 72°C × 2min) × 34, 72°C × 2min
6313p2R	CTTCGTGAAAAGGATTAGAGGT	
2,970 F	ACCAAAAACTGATTTTCAAAGTGG	95°C × 3min, (94°C × 30sec, 54°C × 30sec, 72°C × 30sec) × 34, 72°C × 10min
3,668 R	AGCTATTCCTGAATCTGGTTCA	
3,582 F	GCAGAAAAAGGATATGTTGAAGGA	
4,280 R	CTGCCTGGCTAGTTCAGTCTTA	
4,176 F	TTACATGGGAAAGTACATCGCA	
4,875 R	CGTAGTCATTATCGATTTCCTTGA	
4,805 F	AGAAGAGTACGAAAAGAAAACGC	
5,498 R	TTGGCTTGCTAGTGTAATGGC	
5,381 F	GTACCAAGTTTGTGGTGATCCTC	
6,647 R	CACCTAACGCGGCTTAAGCG	

Supplemental Table 2-2. Primers and thermal cycling conditions used to amplify a region of the *Wolbachia* surface protein gene, *wsp* in order to screen for the presence of *Wolbachia* symbionts.

Region amplified	Primer name	Sequence	PCR conditions	Source
Wolbachia surface protein, *wsp* gene	wspF	TGGTCCAATAAGTGATG AAGAAAC	95°C × 3min, (94°C × 1min, 54°C × 1min, 72°C × 2min) × 34, 72°C × 10min	(Zhou et al. 1998)
	wspR	AAAAATTAAACGCTACT CCA		

Supplemental Table 2-3. Primers and thermal cycling conditions used to amplify across a putatively spliced intron in Odile4 cDNA and gDNA. Primers were designed to bind to the first and second exons of Odile4. Samples were run with two different extension times in order to compare gDNA and cDNA product sizes.

Region amplified	Primer name	Sequence	PCR conditions
Odile4	6316_23 F	ACAAAGACGCTTCTGTTGGAATGC	98°C × 30sec, (98°C × 10 sec, 64°C × 20 sec, 72°C × 10 sec or 45 sec) × 34, 72°C × 2 min
	6316_1905 R	CTCGCACCTTCTCGGGACG	

Supplemental Table 2-4. Nested primers and thermal cycling conditions used to amplify across a putative intron in Odile4. PCR product generated with primers from Table S2-3 from gDNA in distorter *Liposcelis* sp.

Region amplified	Primer name	Sequence	PCR conditions
Odile4	6316_47F	AATCACGCAATAACTGGCTGCAAGC	94°C × 30sec, (94°C × 10 sec, 58°C × 20 sec, 72°C × 50 sec) × 34, 72°C × 2 min
	6316nes 1 R	CTTTCCTACATCGATAAGTACAGAC	
	6316nest2 F	CATTGTGTACTGCAGCGATTAGAG	
	6316nes2 R	TCTGCCGGTAAGTTTACCATGG	
	6316nest3 F	CCGATCCCTAGAACCCAGC	
	6316_1712 R	GATTGAATTTCCCTATAGTCCTGTTC	

Supplemental Table 2-5. GenBank accession numbers of amino acid sequences used in the phylogenetic analysis of the Odile gene in *Liposcelis* sp..

Individual in phylogeny	Accession Numbers
Candidatus Paracaedibacter acanthamoebae	WP_038464660.1
Rhagoletis zephyria	XP_017472977.1
Wasmannia auropunctata	XP_011686120.1
Wolbachia endosymbiont of *Culex quinquefasciatus* (wPip)	WP_007302887.1
Wolbachia endosymbiont of *Aedes albopictus* (wAlb)	PCG12683.1
Wolbachia endosymbiont of *Drosophila ananassae* (wAna)	EAL58666.1
Wolbachia endosymbiont of *Drosophila melanogaster* (wMel)	WP_022626293.1
Wolbachia endosymbiont of *Drosophila recens* (wRec)	WP_080717617.1

Supplemental Table 2-5. *Continued.*

Individual in phylogeny	Accession Numbers
Wolbachia endosymbiont of *Drosophila simulans* (WAu)	CDR78945.1
Wolbachia endosymbiont of *Drosophila simulans* (wNo)	WP_015587726.1
Wolbachia endosymbiont of *Drosophila* sp. (multiple species)	WP_084741728.1
Wolbachia endosymbiont of *Folsomia candida* (wFcan)	APR98581.1
Wolbachia endosymbiont of *Muscidifurax uniraptor* (wUni)	WP_006280034.1
Wolbachia endosymbiont of *Nasonia vitripennis* (wVitA)	WP_082246101.1
Wolbachia endosymbiont of *Nomada* sp. (multiple species)	WP_065094676.1
Wolbachia endosymbiont of *Plutella australiana* (wAus)	PBQ26533.1

Supplemental Table 2-6. Degenerate primers designed for PCR to amplify homologs of Odile1 in *Liposcelis bostrychophila 'AF'*, 'CGC' and *'*granicola*'* for maximum likelihood phylogeny.

Region amplified	Primer name	Sequence	PCR conditions
Odile homolog	HGTdegF	GGAAGAAAGAAWGAAYTACHT GATTT	95°C × 3 min, (95°C × 1 min, 54°C × 30 sec, 72°C × 1 min) × 34, 72°C × 10 min
	HGTdegR	TTCTTTTCGTACTCTTCTAGATAY TTA	

Supplemental Table 2-7. GenBank accession numbers of nucleotide sequences used in the phylogenetic analysis of the Odile gene in *Liposcelis* sp.

Individual in phylogeny	Accession Numbers
Candidatus Paracaedibacter acanthamoebae	CP008941.1
Wasmannia auropunctata	XM_011687818.1
Wolbachia endosymbiont of *Culex quinquefasciatus* (wPip)	AM999887.1
Wolbachia endosymbiont of *Drosophila simulans* (wNo)	CP003883.1
Wolbachia endosymbiont of *Folsomia candida* (wFcan)	CP015510.2
Wolbachia endosymbiont of *Drosophila incompta* (wInc_SM)	CP011149.1
Wolbachia endosymbiont of *Drosophila incompta* (wInc_Cu)	CP011148.1
Wolbachia endosymbiont of *Drosophila melanogaster* (wMel)	AE017196.1
Wolbachia endosymbiont of *Drosophila simulans* (wAu)	LK055284.1
Wolbachia endosymbiont of *Drosophila simulans* (wHa)	CP003884.1
Wolbachia endosymbiont of *Drosophila simulans* (wRi)	CP001391.1

Supplemental Table 2-8. GenBank accession numbers of nucleotide sequences used in the phylogenetic analysis of COI from *Liposcelis* species.

Individual in phylogeny	Accession Numbers
Liposcelis bostrychophila Beibei strain	JN645275.1
Liposcelis bostrychophila CQ isolate	HQ01887.1
Liposcelis bostrychophila CB isolate	KJ680307.1
Liposcelis corrodens CZ isolate	HQ658139.1
Liposcelis sculptilimacula	KX171073.1
Liposcelis paeta	KF649226.1
Liposcelis entomophila	KF649223.1
Liposcelis decolor	JX870621.1
Liposcelis sp. distorter	KP657699.1
Liposcelis sp. nondistorter	KP657693.1
Psococerastis albimaculata	JQ910989.1

Supplemental Table 2-9. DNA primers used to amplify and sequence COI mitochondrial gene in *Liposcelis* sp., *L. bostrychophila 'AF'*, 'CGC', 'MB' and *'*granicola*'* for maximum likelihood phylogeny.

Region amplified	Primer name	Sequence	PCR conditions	Source
Cytochrome c oxidase subunit I	LCO1490	GGTCAACAAATCATAAAGATATTGG	95°C × 3 min, (95°C × 1 min, 54°C × 30 sec, 72°C × 1 min) × 34, 72°C × 10 min	(Folmer et al. 1994)
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA		

Supplemental Table 2-10. Primers and qPCR cycling conditions used for RT-QPCR analysis of Odile1.

Region amplified	Primer name	Sequence	PCR conditions
Odile1	DF6313 F	TTCAAGAACATGGTGAAGACTTCC	95°C × 10min, (95°C × 5 sec, 60°C × 30 sec) × 39, 95°C × 10sec, 65˚C-95˚C increasing 0.5˚C/cycle
	DF6313 R	TCTTTTCCATTTCCATCTTTTGTTTCA	
Ribosomal protein P0	RPLP0 F	CGCATGTCTCTTAGAGGTACAG	
	RPLP0 R	TTTCAAGAGCCTGATTTCGTTC	
Glyceraldehyde 3-phosphate dehydrogenase	GAPDH F	CCAGCTACGATGAAATCAAAGC	
	GAPDHR	CCAGCTACGATGAAATCAAAGC	
Telomeric repeat binding factor	Trf2 F	TCTTGGCTGCAAGAAAATATGC	
	Trf2 R	ACATCACAGGTCCCAACTATGT	
Elongation factor 1-alpha	Ef1a F	TCAAGAAGATCGGTTACAACCC	
	Ef1a R	TCCTTCCGTTCAATCTTCCATC	
Ribosomal protein L32	RPL32 F	GCACAAAGCGTTTCATAAGG	
	RPL32 R	AGCATATGTCGTGTACGTT	

Chapter 2 – References

Armenteros, J.J.A., Sønderby, C.A., Sønderby, S.K., Nielsen, H., and Winther, O. 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics **21**(33): 3387–3395.

Beiko, R.G., Harlow, T.J., and Ragan, M.A. 2005. Highways of gene sharing in prokaryotes. Proc. Natl. Acad. Sci. **102**(40): 14332–14337. doi:10.1073/pnas.0504068102.

Bhasin, M., and Raghave, G.P.S. 2010. ESLpred: SVM-based method for subcellular localization of eukaryotic proteins using dipeptide composition and PSI-BLAST. Nucleic Acids Res. **32**(Web server issue): W414–W419.

van der Biezen, E.A., and Jones, J.D.G. 1998. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr. Biol. **8**(7): R226–R228. doi:10.1016/S0960-9822(98)70145-9.

Blatch, G.L., and Lässle, M. 1999. The tetratricopeptide repeat: A structural motif mediating protein-protein interactions. BioEssays **21**(11): 932–939. doi:10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N.

Boomsma, J.J. 1991. Adaptive colony sex ratios in primitively eusocial bees. Trends Ecol. Evol. **6**(3): 92–95. doi:10.1016/0169-5347(91)90182-W.

Burt, A., and Trivers, R. 2006. Genes in Conflict: The Biology of Selfish Genetic Elements. Belknap Press, Cambridge, MA.

Craig, J.P., Bekal, S., Hudson, M., Domier, L., Niblack, T., and Lambert, K.N. 2008. Analysis of a horizontally transferred pathway involved in vitamin B6 biosynthesis from the soybean cyst nematode Heterodera glycines. Mol. Biol. Evol. **25**(10): 2085–2098. doi:10.1093/molbev/msn141.

Criscuolo, A., and Gribaldo, S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. **10**(210).

Danchin, E.G.J., Rosso, M.-N., Vieira, P., de Almeida-Engler, J., Coutinho, P.M., Henrissat, B., and Abad, P. 2010. Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc. Natl. Acad. Sci. U. S. A. **107**(41): 17651–17656. doi:10.1073/pnas.

Das, A.K., Cohen, P.W., Barford, D., Abraham, J., Leslie, A., Aitken, A., . . . Gamblin, S. 1998. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. EMBO J. **17**(5): 1192–9. doi:10.1093/emboj/17.5.1192.

Doolittle, W.F., and Sapienza, C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature **284**(5757): 601–603. doi:10.1038/284601a0.

Dunning-Hotopp, J.C., Clark, M.E., Oliveira, D.C.S.G., Foster, J.M., Fischer, P., Munoz, M.C., Giebel, J.D., Kumar, N., Ishmael, N., Wang, S., Ingram, J., Nene, R. V., Shepard, J., Tomkins, J., Richards, S., Spiro, D.J., Ghedin, E., Slatko, B.E., Tettelin, H., and Werren, J.H. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science. **317**(5845): 1753–1756.

Dunning Hotopp, J.C. 2011. Horizontal gene transfer between bacteria and animals. Trends Genet. **27**(4): 157–163. doi:10.1016/j.tig.2011.01.005.

Engelstädter, J., and Hurst, G.D.D. 2009. The ecology and evolution of microbes that manipulate host reproduction. Annu. Rev. Ecol. Evol. Syst. **40**: 127–149. doi:10.1146/annurev.ecolsys.l.

Feng, S., Yang, Q., Li, H., Song, F., Stejskal, V., Opit, G.P., Cai, W., Li, Z., and Shao, R. 2018. The highly divergent mitochondrial genomes indicate that the booklouse, Liposcelis bostrychophila (Psocoptera: Liposcelididae) is a cryptic species. G3 genes, genomes, Genet. **8**(3): 1039–1047. doi:10.1534/g3.117.300410.

Fisher, R.A. 1930. The Genetical Theory Of Natural Selection. Clarendon Press, Oxford, England.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. **3**(5): 294–299.

Fournier, D., Estoup, A., Orivel, J., Foucaud, J., Jourdan, H., Le Breton, J., and Keller, L. 2005. Clonal reproduction by males and females in the little fire ant. Nature **435**(7046): 1230–1234. doi:10.1038/nature03705.

Gladyshev, E. a, Meselson, M., and Arkhipova, I.R. 2008. Massive horizontal gene transfer in bdelloid rotifers. Science. **320**(5880): 1210–1213. doi:10.1126/science.1156407.

Gouy, M., Guindon, S., and Gascuel, O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. **27**(2): 221–224.

Groves, M.R., and Barford, D. 1999. Topological characteristics of helical repeat proteins. Curr. Opin. Struct. Biol. **9**(3): 383–389. doi:10.1016/S0959-440X(99)80052-9.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., and Yeats, C. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. **59**(3): 307–321.

Haig, D. 1993. The evolution of unusual chromosomal systems in coccoids: extraordinary sex ratios revisited. J. Evol. Biol. **6**(1): 69–77.

Hamilton, P.T., Hodson, C.N., Curtis, C.I., and Perlman, S.J. 2018. Genetics and genomics of an unusual sex ratio distortion in an insect. Curr. Biol. **28**(23): 3864-3870.

Heimpel, G.E., and de Boer, J.G. 2008. Sex Determination in the Hymenoptera. Annu. Rev. Entomol. **53**(1): 209–230. doi:10.1146/annurev.ento.53.103106.093441.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., and Werren, J.H. 2008. How many species are infected with *Wolbachia*? - A statistical analysis of current data. FEMS Microbiol. Lett. **281**(2): 215–220. doi:10.1111/j.1574-6968.2008.01110.x.

Hodson, C.N., Hamilton, P.T., Dilworth, D., Nelson, C.J., Curtis, C.I., and Perlman, S.J. 2017. Paternal Genome Elimination in *Liposcelis* Booklice. Genetics **206**(June): 1091–1100. doi:10.1534/genetics.117.199786/-/DC1.1.

Hunter, M.S., Perlman, S.J., and Kelly, S.E. 2003. A bacterial symbiont in the Bacteroidetes induces cytoplasmic incompatibility in the parasitoid wasp *Encarsia pergandiella*. Proc. R. Soc. B Biol. Sci. **270**(1529): 2185–2190. doi:10.1098/rspb.2003.2475.

Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., and Yeats, C. 2009. InterPro: the integrative protein signature database. Nucleic Acids Res. **37**(Database issue): D211–D215.

Hurst, G.D.D., and Werren, J.H. 2001. The role of selfish genetic elements in eukaryote evolution. Nat. Rev. Genet. **2**(8): 597–606.

Imai, K., Asakawa, N., Tsuji, T., Akazawa, F., Ino, A., Sonoyama, M., and Mitaku, S. 2008. SOSUI-GramN: high performance prediction for subcellular localization of proteins in Gram-negative bacteria. Bioinformatics **2**(9): 417–421.

Ioannidis, P., Johnston, K.L., Riley, D.R., Kumar, N., White, J.R., Olarte, K.T., Ott, S., Tallon, L.J., Foster, J.M., Taylor, M.J., and Hotopp, J.C.D. 2013. Extensively duplicated and transcriptionally active recent lateral gene transfer from a bacterial *Wolbachia* endosymbiont to its host filarial nematode *Brugia malayi*. : 20–30.

Izard, J.W., and Kendall, D.A. 1994. Signal peptides: exquisitely designed transport promoters. Mol. Microbiol. **13**(5): 765–773. doi:10.1111/j.1365-2958.1994.tb00469.x.

Jaenike, J. 2001. Sex chromosome meiotic drive. Annu. Rev. Ecol. Seystematics **32**(2001): 25–49.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305**(3): 567–580.

Leclercq, S., Thézé, J., Amine, M., Giraud, I., Moumen, B., Ernenwein, L., Greve, P., Gilbert, C., and Cordaux, R. 2016. Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome. Proc. Natl. Acad. Sci. U. S. A. **113**(52): 15036–15041. doi:10.1073/pnas.1608979113.

Lynch, M. 2006. Streamlining and Simplification of Microbial Genome Architecture. Annu. Rev. Microbiol. **60**(1): 327–349. doi:10.1146/annurev.micro.60.080805.142300.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., C.J., L., Lu, S., Farideh, C., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., and Bryant, S.H. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. **45**(D1): D200–D203.

McNulty, S.N., Foster, J.M., Mitreva, M., Hotopp, J.C.D., Martin, J., Fischer, K., Wu, B., Davis, P.J., Kumar, S., Brattig, N.W., Slatko, B.E., Weil, G.J., and Fischer, P.U. 2010. Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. PLoS One **5**(6): e11029. doi:10.1371/journal.pone.0011029.

Mueller, U.G. 1991. Haplodiploidy and the evolution of facultative sex ratios in a primitively eusocial bee. Science. **254**(5030): 442–444.

Nur, U.Z.I., Werren, J.H., Eickbush, D.G., Burke, W.M.D., and Eickbush, T.H. 1988. A "Selfish" B chromosome that enhances its transmission by eliminating the paternal Genome. Science. **240**(4851): 512–514.

de Oliveira, C.D., Gonçalves, D.S., Baton, L.A., Shimabukuro, P.H.F., Carvalho, F.D., and Moreira, L.A. 2015. Broader prevalence of Wolbachia in insects including potential human disease vectors. Bull. Entomol. Res. **105**(3): 305–315. doi:10.1017/S0007485315000085.

van Ooijen, G., Mayr, G., Kasiem, M.M.A., Albrecht, M., Cornelissen, B.J.C., and Takken, F.L.W. 2008. Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. J. Exp. Bot. **59**(6): 1383–1397. doi:10.1093/jxb/ern045.

Perlman, S.J., Hodson, C.N., Hamilton, P.T., Opit, G.P., and Gowen, B.E. 2015. Maternal transmission, sex ratio distortion, and mitochondria. PNAS. **112**(33): 10162–10168. doi:10.1073/pnas.1421391112.

Peterson, T.N., Brunak, S., von Heijne, G., and Nielson, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods **8**: 785–786.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29**(9): 45e–45. doi:10.1093/nar/29.9.e45.

Pond, S.L.K., Frost, D.W., and Muse, S.V. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics **21**(5): 676–679.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available from https://www.r-project.org/.

Rambaut, A. 2007. FigTree v.1.4.2 ed. Available from http://tree.bio.ed.ac.uk/software/figtree/.

Rogozin, I.B., Wolf, Y.I., Sorokin, A. V., Mirkin, B.G., and Koonin, E. V. 2003. Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in Eukaryotic evolution. Curr. Biol. **13**(17): 1512–1517. doi:10.1016/S.

Savojardo, C., Martelli, P.L., Fariselli, P., Profiti, G., and Casadio, R. 2018. BUSCA: an integrative web server to predict subcellular localization of proteins. Nucleic Acids Res. **46**(W1): W459–W466.

Shires, S.W. 1982. Observation on Liposcelis bostrychophilus Badonnel (Psocoptera). Trop. Prod. Inst. **44**: 11–14.

The Nasonia Working Group. 2010. Functional and evolutionary insights from the genomes of three parasitoid Nasonia species. Science. **327**(5963): 343–348.

Werren, J.H. 1997. Biology of *Wolbachia*. Annu. Rev. Entomol. **124**(42): 587–609. doi:doi:10.1146/annurev.ento.42.1.587.

Werren, J.H. 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc. Natl. Acad. Sci. U. S. A. **108**(2): 10863–10870. doi:10.1073/pnas.

Werren, J.H., Baldo, L., and Clark, M.E. 2008. dont use. Nat. Rev Microbiol **6**(10): 741–751. doi:10.1038/nrmicro1969.

Werren, J.H., Nur, U., and Wu, C.I. 1988. Selfish genetic elements. Trends Ecol. Evol. **3**(11): 297–302. doi:10.1016/0169-5347(88)90105-X.

Yu, C.S., Chen, Y.C., Lu, C.H., and Hwang, J.K. 2006. Prediction of protein subcellular localization. Proteins Struct. Funct. Bioinforma. **64**: 643–651.

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., and Brinkman, F.S. 2010. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics **26**(13): 1608–1615.

Zhou, W., Rousset, F., and O'Neill, S. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using wsp gene sequences. Proc. R. Soc. B Biol. Sci. **265**(1395): 509–515. doi:10.1098/rspb.1998.0324.

Zug, R., and Hammerstein, P. 2012. Still a host of hosts for *Wolbachia*: Analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PLoS One **7**(6): 7–9. doi:10.1371/journal.pone.0038544.

Chapter 3 – The role of male mating behaviour in the persistence of a selfish, sex ratio distorting element in *Liposcelis* sp.

Introduction

Classical theories on mate choice consider females to be the 'choosier' sex, while males compete with each other for the attention and opportunity to mate with a female (Darwin, 1874; Bateman, 1948). It is generally thought that females discriminate more often between prospective mating partners as they have to make a greater physiological and metabolic investment per offspring than males (Thornhill, 1976; Trivers, 1972). In sexually reproducing species, the two sexes have vastly different reproductive strategies. Females produce a few 'costly' gametes, while males produce a large number of small 'cheap' gametes in excess (Parker et al. 1972). Because of these differing strategies, there is increased competition for male gametes to fertilize the rarer female gametes, thus subjecting males to stronger sexual selection than females. The choosier sex, in this case females, can therefore select mates based on differences in fitness (i.e. viability, mating success, fecundity) in order to more successfully pass on their genes to the next generation (Emlen and Oring, 1977).

However, mate choice is not only the realm of females. Male mate choice can also evolve when it is beneficial for a male to choose one female over another. In order for male mate choice to evolve, there must be variation in female quality (Edward and Chapman, 2011), for example, when there is a phenotypic indicator of female fecundity, such as body size or the number of mature eggs being carried (Bonduriansky, 2001). Additionally, age, sexual experience, or disease, may have an effect on male mate choice for high-value females (Beltran-Bech and Richard, 2014; Fortin et al. 2018; McDonald and Borden, 1996). In these cases, males may choose to mate with females based on these phenotypic indicators, or invest varying resources in copulation based on mate quality (Otronen, 1984; Wedell and Arak, 1989).

Certain selfish genetic elements that increase their own transmission can impose fitness costs, and if linked to a phenotypic indicator, can be detected by a prospective mate (Lande and Wilkinson, 1999). The genetic conflict that occurs between selfish

genetic elements and their associated genome can occur within an individual's genome (such as selfish X chromosomes), as well as between cytoplasmic elements and nuclear genes (such as endosymbionts). When genetic conflict occurs over a reproductive strategy, this can then result in a sex ratio distortion (Burt and Trivers, 2006). *Wolbachia*, a widespread bacterial endosymbiont of arthropods, is a well-studied example of a selfish cytoplasmic element that has evolved the ability to manipulate its host's reproduction in interesting ways. *Wolbachia* is transmitted maternally through the egg cytoplasm; therefore, by increasing the number of female offspring of their hosts, it can increase its own transmission to the next generation (Hurst et al. 1997). *Wolbachia* has been shown to distort the sex ratio in favour of females through killing or feminizing males (Werren et al. 1986; Rigaud 1997), or inducing parthenogenesis (Stouthamer et al. 1993). This reproductive manipulation can result in an extremely female-biased sex ratio, such that sexual selection may be modified to reverse the sex that competes for a mate, causing males to become the choosier sex (Moreau and Rigaud, 2000). Male-killing *Wolbachia* present in two butterfly species (*Acraea encedon* and *Acraea encedana*) has caused a sex-role-reversed mating system in which females, rather than males, form a lekking swarm to attract the rarer males (Jiggins et al. 2000). As well, in the terrestrial isopod *Armadillidium vulgare*, males discriminate and prefer uninfected females over *Wolbachia*-feminized males (Bouchon et al. 1998; Moreau et al. 2001).

A more extreme example of a reproductive system in which male mate choice has been demonstrated is in gynogenetic species. In these species, females of hybrid origin (gynogens), require sperm from males of a related species to reproduce, but the genetic material from the male is not passed on to the offspring (Vrijenhoek, 1994). These so-called 'sperm parasites' have been documented in 24 genera across seven phyla (Beukeboom and Vrijenhoek, 1998) and have been best studied in the freshwater fish, the Amazon molly (*Poecilia formosa)* (Hubbs and Hubbs, 1932). In order to stimulate embryogenesis, the Amazon molly has to mate with males from either parental species (*P. latipinna* or *P. mexicana)* (Woodhead and Armstrong, 1985). Both *P. latipinna* and *P. mexicana* males have shown a significant preference to mate with conspecifics, rather than with the mollies (Woodhead and Armstrong, 1985; Schlupp, 1991; Ryan et al. 1996). Male mate choice against gynogenetic sperm parasites has also been found in the

bark beetle (*Ips accuminatus*) (Loyning and Kirkendall, 1996) and the Japanese crucian carp (*Carassius auratus langsdorfii*) (Hakoyama and Iguchi, 2002).

In hybridogenesis, a similar mode of reproduction to gynogenesis, hybrids must mate with one of the parental species in order to reproduce (Schultz, 1969). In early development, paternal chromosomes are lost from the germline of the hybrid. Following fertilization with the parental species whose chromosomes were lost, the chromosome set is restored once again to the diploid state (Beukeboom and Vrijenhoek, 1998). Best studied in aquatic frogs (*Rana esculenta*) (Graf and Polls-Pelaz, 1989), hybridogenesis has also been reported in stick insects (*Bacillus rossius-grandii*) (Mantovani and Scali, 1991), and fish in the genera *Poeciliposis* and *Tropido phoxinellus* (Carmona 1997; Schultz 1969). In mate choice trials by Mckay (1971), male *Poeciliposis* chose to mate exclusively with conspecifics; however, hybridogens were also inseminated once male competition was introduced.

Males waste their sperm when they mate with sperm parasites, so it is evolutionarily beneficial for males to avoid wasting resources on these females if they are able to distinguish between the different female types. In these systems, a co-evolutionary arms race may develop as males evolve to avoid mating with the sperm-parasites, and the sperm-parasites evolve to increase mating chances with males. Such populations are predicted to go extinct, with a shortage of males and an excess of sperm-parasites that rely on sperm to reproduce (Heubel et al. 2009). Despite this, populations of sperm-parasites persist, and differences in male mate preference have been proposed as a possible stabilizing factor (Uzzell, 1964).

An extreme female-biased sex ratio distortion has been found in a recently discovered species of booklouse, *Liposcelis* sp. (Order: Psocodea) (Perlman et al. 2015). This sexual species has two distinct female types. Distorter females carry a maternally transmitted selfish genetic element, the distorting element, that results in the production of exclusively female offspring, while the nondistorter females produce both sons and daughters. Both female types require a male to reproduce, so distorter females must mate with the sons of nondistorter females. However, distorter females are not gynogenetic as paternal alleles are found in the daughters of the distorter females. This mode of sex determination is most similar to that of hybridogenesis, as the daughters of distorter

females do not pass on the paternal genes to their offspring (Hamilton et al. 2018). Because the paternal DNA is not passed on to the F2 generation, the act of mating with a distorter female is essentially a 'dead end' for paternal genes, potentially placing a selective pressure on the male to evolve a preference to mate with nondistorter females over the distorter sperm parasites. This could especially be true in this species as males transfer sperm via a spermatophore and mating time is long, suggesting an increased investment in gametes and time. Additionally, the distorter-specific genome is diverging from that of the nondistorter genome, as it does not recombine with the nondistorter genome. If any of the accumulated changes in the distorter genome result in a phenotypic indicator detectable by a male, the development of male mate choice against the distorting element may evolve. Currently it is not known whether male *Liposcelis* sp. can distinguish between the two female types; however, conditions that have previously lead to the development of male mate choice in other populations appear to be present here.

In this chapter, I examine whether male mate choice is present in *Liposcelis* sp., and whether there is expression of paternal genes in distorters that may cover distorter specific signals. I carried out a series of mate choice trials in which individual males were given a choice between nondistorter and distorter females. I recorded which female type mated in each trial, as well as the time to initiate mating and the duration of mating in a subset of trials. I also looked into whether both the maternal and paternal genes were being expressed in the distorter offspring. My aim was to assess whether male mate choice has evolved in *Liposcelis* sp., and to investigate whether male mate choice could be a contributing factor to the persistence of the distorting element. It is still unclear how this distorting element is able to persist in the population without increasing in frequency to a point of instability due to a shortage of males. Little is known about the mating dynamics of sexual *Liposcelis*, especially one that contains a maternally transmitted, chromosomal sex ratio distorter. I hope to provide more information on the mating habits of this non-model insect group, specifically in the role male mate preference may play in the persistence of this sex ratio distortion in the species.

Methods

Booklouse colony information

The *Liposcelis* sp. used in this study were collected in 2010 from the Chiricahua Mountains in southeastern Arizona. I maintain the two female types as separate cultures in the laboratory and keep them at 27°C and 75% relative humidity. The cultures are kept in 125ml glass canning jars that contain a 1:10 (by weight) mixture of Rice Krispies (Kellogg's) to cracked wheat (Bob's Red Mill). The lids of the jars are replaced with a 70mm Whatman filter paper (Sigma-Aldrich). I maintain colonies every 14 days in which food is replaced as needed and males (from the nondistorter colonies) are added to the distorter colonies.

Mating trials

This experiment was set up in order to examine if male *Liposcelis* sp. showed a mating preference toward either one of the two female types. I obtained individuals before they reached adulthood, to ensure all subjects were unmated, and kept them in small petri dishes (35mm in diameter) with 0.7g of food. Because wildtype nymphs can mature into either males or females, I checked these dishes regularly in order to separate the males and females based on morphology, yet still prior to sexual maturation. The three dishes containing the virgin distorter females, and virgin male and nondistorter female booklice were maintained at regular conditions until they reached sexual maturity.

I carried out six separate series of mating trials from October 2017 to March 2018. Each mating trial was set up in a small petri dish (35mm in diameter), where in each I placed one distorter female, one nondistorter female, and one male in each. As soon as mating was observed to have initiated, I removed the female that was not mating. Successful mating was easy to observe, as the male and female stay engaged in the mating process through the formation of a spermatophore (Figure S3-1) for a long period of time (average mating time was 70 minutes). I timed thirty-two of the successful

matings and recorded both the time to initiate mating, and the duration of mating (Table S3-1).

I extracted DNA from both the mated and un-mated females in order to genotype. I extracted DNA using 30uL PrepMan Ultra (Life Technologies) to obtain 15uL of DNA. In order to distinguish between distorter and nondistorter females, I used a primer set that targets a gene present only in distorters (Odile1) (Table S3-2). I visualized amplicons on an agarose gel and determined female genotype for each sample based on presence/absence of a band. I genotyped un-mated females in order to confirm that the males were indeed choosing between the two different female types. I analyzed data in R Studio v1.0.136 (R Core Team, 2016).

Paternal Gene Expression

In order to determine whether paternal genes were being expressed in distorters, I sequenced cDNA from the cAMP-specificIBMX- insensitive 39,59-cyclic phosphodiesterase gene (phos1) to look for heterozygosity. Variation at this gene was previously used to track maternal and paternal inheritance in *Liposcelis* sp. (Perlman et al. 2015; Hodson et al. 2017; Hamilton et al. 2018). I extracted total RNA from five distorter females, five nondistorter females and four males using Trizol-LS (Invitrogen). I treated 10 uL of the RNA from each individual with DNaseI endonuclease (Thermoscientific) to remove any remaining genomic DNA. I reverse transcribed the RNA to cDNA using SuperScriptII (Invitrogen) and with the use of Oligo dts(20) (IDT) to target the poly-A tail of the mRNA transcripts. I used phos1 specific primers (Table S3-3) in a PCR and confirmed amplification on an agarose gel. Phos1 amplicons were sequenced (Sequetech) and I looked for heterozygosity by examining the chromatograms for double peaks (Geneious 7.1.9) (Figure S3-2).

Results

Mate choice

There was no significant difference between which female type, distorter or nondistorter, the male *Liposcelis* sp. mated with (X^2 (1, N = 66) = 1.5, p = .22) (Table 3-1). Of the 66 successful matings where the male had a choice between the two female types, the male chose to mate with distorter females 28 times and nondistorter females 38 times. Genotyping confirmed that each mating trial contained one distorter female and one nondistorter female.

Table 3-1. Male mate choice in mating trials between distorter and nondistorter female *Liposcelis* sp. Successful mating occurred in 66 trials, from which DNA was extracted from the females and subsequently genotyped with PCR using a distorter specific primer set. A Chi-square test of independence was used in order to determine whether there was a significant difference between which female was chosen more often.

Female type	Mated	Chi-square test of independence
Distorter	28	X^2 (1) = 1.5 p = 0.22 n = 66
nondistorter	38	

Timed mating trials

I used a one-way ANOVA to compare the time it took to initiate mating to the two female types of *Liposcelis* sp., distorter and nondistorter. There was not a significant difference between the two female types (distorter or nondistorter) in the time it took for mating to begin between the male and one of the females [F(1) = 1.65, p = 0.21] (Figure 3-1 A). Additionally, I did not see a significant difference between the two female types and how long mating lasted for [F(1) = 0.65, p = 0.443] (Figure 3-1 B).

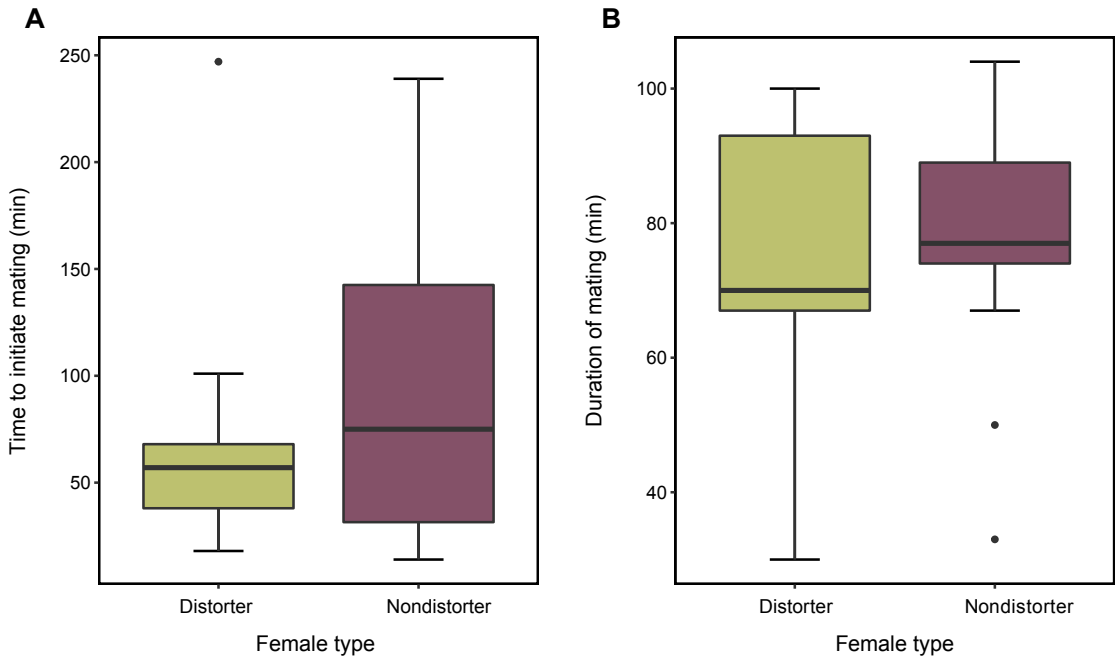


Figure 3-1. Timed mating trials of distorter and nondistorter *Liposcelis* sp. **A.** Time (min) from the addition of the male *Liposcelis* sp. until mating began for distorter (n=28) and nondistorter (n=38) *Liposcelis* sp. **B.** Duration (min) of mating for distorter (n=28) and nondistorter (n=38) *Liposcelis* sp. Time for mating to initiate did not significantly differ between the two female types, nor did the duration of mating.

Because the timed mating trials occurred on three different dates, I checked to see if trial number had an effect on the time to initiate mating or the duration of mating between the two genotypes. I used a two-way ANOVA to compare the effect of genotype and trial on initiation time and found that there was no significant difference [$F(3) = 0.734$, $p = 0.54$]. Additionally, I used a two-way ANOVA to compare the effect of genotype and trial on the duration of mating and found no significant difference [$F(3) = 0.572$, $p = 0.64$].

Paternal gene expression

I found that the distorter females (n = 5) were heterozygous at the three loci examined on the phos1 gene. An allele from the father and an allele from the distorter mother were both present in the chromatograms generated from Sanger sequencing of

phos1 cDNA (Table 3-2). Additionally, I looked at chromatograms from male cDNA (n = 4) and nondistorter cDNA (n=5) at the same three loci of the phos1 gene and all were homozygous.

Table 3-2. Allelic expression of the phos1 gene. Nucleotides at three SNPs were determined by examining chromatograms generated by Sanger sequencing of cDNA from distorter (DF) (n=5), male (M) (n=4) and nondistorter (N) (n=5) *Liposcelis* sp. Distorters were heterozygous while males and nondistorter females were homozygous at the three loci of phos1 that were examined.

Individual	SNP 1	SNP 2	SNP 3
DF1	GA	TC	GC
DF2	GA	TC	GC
DF3	GA	TC	GC
DF4	GA	TC	GC
DF5	GA	TC	GC
M1	AA	CC	CC
M2	AA	CC	CC
M3	AA	CC	CC
M4	AA	CC	CC
N1	AA	CC	CC
N2	AA	CC	CC
N3	AA	CC	CC
N4	AA	CC	CC
N5	AA	CC	CC

Discussion

Information on female mate choice is well-known across many taxa, and increasing evidence for the presence of widespread mate choice in males is becoming apparent as well. However, little is known about the role male mate choice may play in the insect Order Psocodea, specifically in the booklouse, *Liposcelis* sp. I examined whether male mate choice may have evolved in this species for defense against the presence of a selfish genetic element that results in 100% female progeny. Previous studies have shown male mate choice has evolved in similar systems where it would be beneficial for a male to distinguish between potential mates, but male *Liposcelis* sp. did not show evidence of making a choice. When male booklice were given the choice between mating with either a nondistorter female, or a distorter female carrying a distorting element, there was no significant difference between which female was chosen. Additionally, time to initiate mating and duration of mating did not significantly differ between the two female types. This apparent lack of discrimination by the males to distinguish between the non-distorter females and the distorter sperm parasites may contribute to the persistence of the selfish genetic element in the population.

The offspring of a male and a distorter female will only ever pass on the maternal distorter genes inherited from the mother resulting in an evolutionary 'dead end' for the paternal genes. Theoretically, selection would favour males that avoid mating with distorter females as they will more successfully pass on their genes to future generations in doing so. However, in order for mate choice to evolve, variation must exist in such a way that is detectable by the choosing sex. Male *Liposcelis* sp. did not exhibit a choice between either female type which suggests that they do not detect a phenotypic indicator of fitness. Chromosomal selfish elements may be detected when they are linked to a detectable allele at a nearby trait locus (Lande and Wilkinson, 1999); however, in the case of *Liposcelis* sp., the phenotypic indicator does not have to be linked to the selfish element as the distorter and nondistorter populations do not recombine.

Although not an example of male mate choice but female mate choice, a well-studied example of phenotypic linkage to a selfish genetic element occurs in two species of stalk-eyed flies (*Teleopsis dalmanni* and *T. whitei*). Males that carry a driving X chromosome (or driving X) typically produce highly female biased broods of 95% or

more females as a result of Y-bearing spermatid degeneration (Presgraves et al. 1997; Wilkinson et al. 2006). The driving X was found to be linked to a shorter eye-stalk span in males allowing females to distinguish between males of high genetic quality and those that carry the driving X (Wilkinson et al. 1998; Johns et al. 2005). Females that avoid mating with males carrying the driving X will produce more sons who will have increased reproductive success in the female-biased population (Lande and Wilkinson, 1999). Previous work has shown that females choose between males based on eye-stalk length and will more often mate with males that have a longer eye-stalk length than those with shorter eye-stalks (and a driving X) (Burkhardt and Ingrid, 1988; Hingle et al. 2001; Wilkinson and Reillo, 1994; Wilkinson et al. 1998). Unlike the stalk-eyed flies, the two female types of *Liposcelis* sp. are indistinguishable to the human eye; however, indicators of selfish genetic elements are not always as pronounced as what is seen with the varying lengths of eye-stalks. Male isopods (*A. vulgare)* were found to distinguish between feminized males (infected with feminizing *Wolbachia*) and uninfected, true females based on odour cues (Fortin et al. 2018).

The distorter-specific portion of the genome differs from that of the nondistorter genome as it does not recombine with nondistorters. This presents the opportunity for distorter-specific phenotypic signals to evolve and be recognized by the males. As of yet, there is no evidence that males can detect any differences between the two females. A few possibilities for this may be that there are insufficient differences between the two genomes, differences do not result in a phenotypic indicator, or differences can be detected but are not important to the males. Even though the paternal genes are not passed on to the next generation in distorters, I found that they are transcribed, at least in part. If the paternal genes are being expressed in the distorter females, this may mask distorter-specific signals and render differentiation by the males impossible.

Several other studies report the absence of male discrimination against females that carry a selfish, sex ratio distorting element. In *Drosophila simulans*, males infected with *Wolbachia* must mate with infected females in order to produce viable offspring (Hoffmann and Turelli, 1988). When *Wolbachia*-infected male *D. simulans* were given a choice between infected or uninfected females, Champion De Crespigny and Wedell (2007) found no evidence of male mate preference. It was proposed that there were no

physiological changes associated with *Wolbachia* infection making it difficult for the males to avoid the less desirable mates. Similarly, no male mate choice was observed in *Drosophila innubila* against females infected with a male-killing *Wolbachia* in which all combinations of virgin and non-virgin males and females were tested (Sullivan and Jaenike, 2006). Another study hypothesized that *Wolbachia* infection would increase the attractiveness of its *D. melanogaster* host in order to further its own spread, but a preference for (or against) infected females was not found (Arbuthnott et al. 2016).

Although no evidence of pre-copulatory mate preference was observed in this study, the possibility of post-copulatory mate preference cannot be ruled out. During copulation, male *Liposcelis* sp. transfer sperm to females as a spermatophore which protects sperm until the transfer has been completed. Sperm transferred to the females are maintained in a spermatheca that stores sperm until they are released for fertilization (Finlayson, 1948). The resources required to produce a spermatophore can be a limiting factor to male fitness, and can therefore result in sperm allocation based on female quality (Wedell et al. 2002). This differential allocation of resources can be defined as 'cryptic' male mate choice, and can include variation in the size of ejaculate or duration of copulation (Bonduriansky, 2001).

Cryptic male mate choice has been found in a number of species and has been shown to play a role in the persistence of selfish elements and sex ratio distorters within populations. Sperm depletion was found to limit fertility of *Wolbachia* infected female isopods (*A. vulgare*) and was hypothesized to help explain the maintenance of the infection in the population without it spreading to fixation (Rigaud and Moreau, 2004). Additionally, differential sperm allocation has been found in gynogenetic species in which males otherwise waste their sperm when mating with asexual individuals (Schlupp and Plath, 2005). Although sperm quantity was not specifically examined in the mating trials of *Liposcelis* sp., the duration of mating could be an indicator of sperm allocation, assuming the duration of sperm transfer correlates to the quantity of sperm transferred to the female. As there was no significant difference found in the time males mated with either female type, it is difficult to conclude whether cryptic male mate choice is present in *Liposcelis* sp.

The age of the sex ratio distorting element in *Liposcelis* sp. is predicted to be ~450-852,000 years (Hamilton et al. 2018). The apparent indifference of the male to choose which female he mates with could provide some insight into how the distorting element has persisted over time within this population. The males mate with distorter females even though they are wasting their sperm in doing so, thus, allowing the distorting element to be passed on to the next generation. Males are limited in the population, perhaps facilitating matings with multiple females, reducing the necessity of having to choose the best mate. If the act of mating is not a significant resource drain on the male, then there would be no apparent benefit to choose one female type over the other.

Larger scale mating trials with different ratios of males to females and the two female types could provide interesting insights into the otherwise unfamiliar mating habits of the male *Liposcelis* sp. and help to further explain the persistence of the distorting element in the population. Virgin males were used in the mating trials to ensure all individuals had the same background of sexual experience; however, there has been evidence of recently mated male *Drosophila melanogaster* exhibiting a greater bias compared to virgin males when choosing a mate (Byrne and Rice, 2006). It would be interesting to see if male mate choice would be similarly affected once mating resources have been depleted in male *Liposcelis* sp. Discrimination between the two female types could have been missed if virgin males mate indiscriminately while non-virgin males are resource-limited and choosier.

Selfish genetic elements that cause sex ratio biases have been known to induce changes in the mating habits of their hosts. *Liposcelis* sp. provides a unique opportunity to examine the persistence of a selfish genetic element causing an extreme sex ratio distortion in an insect population. Little is known about the mating habits of male *Liposcelis* sp., but perhaps their mating behaviours play a role in the persistence of the distortion. Following mate choice trials, I found that male *Liposcelis* sp. did not make an observable choice between distorter or non-distorter females as mating partners. Additionally, neither the time to initiate mating nor the duration of mating differed between the two female types. Perhaps this lack of discrimination plays a role in the persistence of the sex ratio distorting element in the population.

Chapter 3 – Supplemental Information

Supplemental Table 3-1. Individual mating trial results including genotype, time until mating began once the male was added, and the duration of mating.

Female type	Time until mating (min)	Duration of mating (min)
Nondistorter	14	79
Distorter	18	68
Nondistorter	26	91
Nondistorter	29	50
Distorter	57	72
Nondistorter	127	73
Nondistorter	118	75
Nondistorter	232	100
Nondistorter	239	75
Nondistorter	235	77
Distorter	247	48
Nondistorter	228	33
Distorter	35	94
Distorter	35	93
Distorter	57	93
Nondistorter	65	104
Nondistorter	66	76
Nondistorter	80	104
Nondistorter	139	81
Nondistorter	146	80

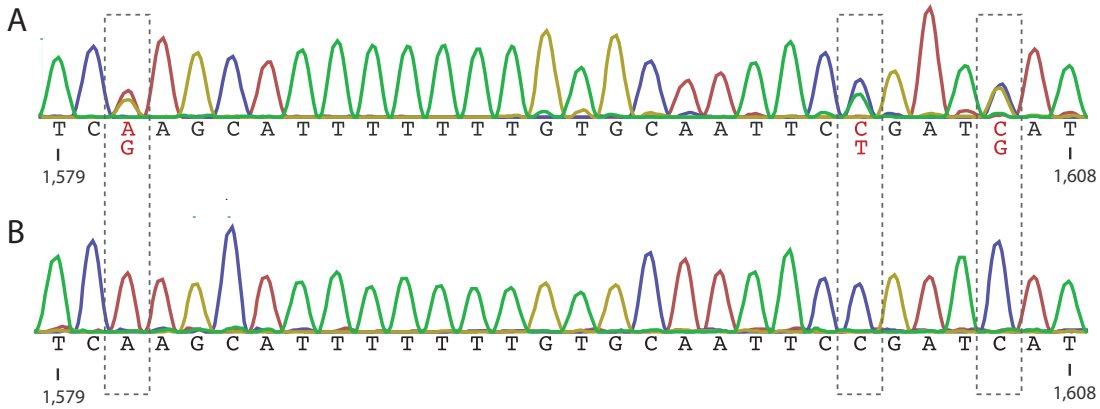
Supplemental Table 3-2. Distorter specific primer sequences and thermocycling conditions used for genotyping *Liposcelis* sp.. Primer sequence includes T7 promoter at 5' end.

Region amplified	Primer name	Sequence	PCR conditions
Odile1	006313_F	TAATACGACTCACTATAGGG AGCGGGTTAGGTGGTTTAGGT	95°C × 3min, (94°C × 1min, 54°C × 1min, 72°C × 1min 15sec) × 34, 72°C × 10min
	006313_R	TAATACGACTCACTATAGGG ATTTGACGAGCAAACGCACAG	

Supplemental Table 3-3. Primers and thermocycling conditions used to amplify a region of the cAMP-specific IBMX-insensitive 39,59-cyclic phosphodiesterase (Phos1) gene used to determine paternal expression in distorters.

Region amplified	Primer name	Sequence	PCR conditions	Source
cAMP-specific IBMX - insensitive 39,59 - cyclic phosphodiesterase	Phos1F	TCCCTTCCGTC AATAAATGC	95°C × 3min, (94°C × 1min, 54°C × 1min, 72°C × 1min 15sec) × 34, 72°C × 10min	Hodson et al. 2017
	Phos1R	AATGTTCGAA ATGCCGAGTC		

Supplemental Figure 3-1. Male (left) and female (right) *Liposcelis* sp. during sperm transfer via spermatophore.



Supplemental Figure 3-2. Chromatogram of Phos1 cDNA segment (Sequetech) from *Liposcelis* sp.. **A.** Phos1 sequence from *Liposcelis* sp. distorters depicting a double peak at the three loci indicated. **B.** Phos1 sequence from *Liposcelis* sp. males depicting a single peak at the same three loci.

Chapter 3 – References

Arbuthnott, D., Levin, T.C., and Promislow, D.E.L. 2016. The impacts of *Wolbachia* and the microbiome on mate choice in *Drosophila melanogaster*. J. Evol. Biol. **29**(2): 461–468. doi:10.1111/jeb.12788.

Bateman, A.J. 1948. Intra-sexual selection in *Drosophila*. Heredity (Edinb). **2**(3): 349–368. doi:10.1038/hdy.1948.21.

Beukeboom, L.W., and Vrijenhoek, R.C. 1998. Evolutionary genetics and ecology of sperm-dependent parthenogenesis. J. Evol. Biol. **11**(6): 755–782.

Bonduriansky, R. 2001. The evolution of male mate choice in insects: A synthesis of ideas and evidence. Biol. Rev. Camb. Philos. Soc. **76**(3): 305–339. doi:10.1017/S1464793101005693.

Bouchon, D., Rigaud, T., and Juchault, P. 1998. Evidence for widespread *Wolbachia* infection in isopod crustaceans: molecular identification and host feminization. Proc. R. Soc. London (March).

Burkhardt, D., and Ingrid, de la M. 1988. Big "antlers" are favoured: female choice in stalk-eyed flies (Diptera, Insecta), field collected harems and laboratory experiments. J. Comp. Physiol. A **162**(5): 649–652.

Carmona, J.A., Sanjur, O.I., Doadrio, I., Machordom, A., and Vrijenhoekt, R.C. 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: The *Tropidophoxinellus alburnoides* complex. Genetics **146**: 983–993.

Darwin, C.R. 1874. The descent of man and selection in relation to sex (2nd ed.). Hurst and Company, New York, NY.

Edward, D.A., and Chapman, T. 2011. The evolution and significance of male mate choice. Trends Ecol. Evol. **26**(12): 647–654. doi:10.1016/j.tree.2011.07.012.

Emlen, S.T., and Oring, L.W. 1977. Ecology, sexual selection, and the evolution of mating systems. Science. **197**(4300): 215–223. doi:10.1126/science.327542.

Finlayson, L.H. 1948. The life-history and anatomy of *Lepinotus patruelis* Pearman(Psocoptera-Atropidae). Proc. Zool. Soc. London **119**(2): 301–323.

Fortin, M., Debenest, C., Souty-Grosset, C., and Richard, F.-J. 2018. Males prefer virgin females, even if parasitized, in the terrestrial isopod *Armadillidium vulgare*. Ecol. Evol. **8**(6): 1–13. doi:10.1002/ece3.3858.

Hakoyama, H., and Iguchi, K. 2002. Male mate choice in the gynogenetic-sexual complex of crucian carp, *Carassius auratus*. Acta Ethol. **4**(2): 85–90. doi:10.1007/s102110100045.

Hamilton, P.T., Hodson, C.N., Curtis, C.I., and Perlman, S.J. 2018. Genetics and genomics of an unusual sex ratio distortion in an insect. Curr. Biol. **28**(23): 3864-3870.

Heubel, K.U., Rankin, D.J., and Kokko, H. 2009. How to go extinct by mating too much : population consequences of male mate choice and efficiency in a sexual-asexual species complex. Oikos **118**(4): 513–520.

Hingle, A., Fowler, K., and Pomiankowski, A. 2001. Size-dependent mate preference in the stalk-eyed fly *Cyrtodiopsis dalmanni*. Anim. Behav. **61**(3): 589–595. doi:10.1006/anbe.2000.1613.

Hodson, C.N., Hamilton, P.T., Dilworth, D., Nelson, C.J., Curtis, C.I., and Perlman, S.J. 2017. Paternal Genome Elimination in *Liposcelis* Booklice. Genetics **206**(June): 1091–1100. doi:10.1534/genetics.117.199786/-/DC1.1.

Hoffmann, a a, and Turelli, M. 1988. Unidirectional incompatibility in Drosophila simulans: inheritance, geographic variation and fitness effects. Genetics **119**(2): 435–444.

Hubbs, C.L., and Hubbs, L.C. 1932. Apparent Parthenogenesis in Nature, in a Form of Fish of Hybrid Origin. Science. **76**(1983): 628–630.

Hurst, G.D.D., Hurst, L.D., and Majerus, M.E.N. 1997. Cytoplasmic Sex Ratio Distorters. *In* Influential Passengers: inherited microorganisms and arthropod reproduction. *Edited by* S.L. O'Neill, A.A. Hoffmann, and J.H. Werren.

Jiggins, F.M., Hurst, G.D.D., and Majerus, M.E.N. 2000. Sex-ratio-distorting Wolbachia causes sex-role reversal in its butterfly host. Proc. R. Soc. B **267**(1438): 69–73.

Johns, P.M., Wolfenbarger, L.L., and Wilkinson, G.S. 2005. Genetic linkage between a sexually selected trait and X chromosome meiotic drive. Proc. R. Soc. B **272**(1576): 2097–2103. doi:10.1098/rspb.2005.3183.

Lande, R., and Wilkinson, G.S. 1999. Models of sex-ratio meiotic drive and sexual selection in stalk-eyed flies. Genet. Res. (Camb). **74**(3): 245–253.

Loyning, M.K., and Kirkendall, L.R. 1996. Mate discrimination in a pseudogamous bark beetle (Coleoptera: Scolytidae): male *Ips acuminatus* prefer sexual to clonal females. Oikos **77**(2): 336–344.

Mantovani, B., and Scali, V. 1991. Hybridogenesis and androgenesis in the stick-Insect *Bacillus rossius- Grandii benazzii* (Insecta, Phasmatodea). Evolution (N. Y). **46**(3): 783–796.

McDonald, R.S., and Borden, J.H. 1996. Courtship behavior and discrimination between potential mates by male *Delia antiqua* (Diptera: Anthomyiidae). J. Insect Behav. **9**(6): 871–885.

Mckay, F.E. 1971. Behavioral aspects of population dynamics in unisexual-bisexual Poeciliopsis (Pisces: Poeciliidae). Ecology **52**(5): 778–790.

Moreau, J., Bertin, A., Caubet, Y., and Rigaud, T. 2001. Sexual selection in an isopod with *Wolbachia*-induced sex reversal: Males prefer real females. J. Evol. Biol. **14**(3): 388–394. doi:10.1046/j.1420-9101.2001.00292.x.

Moreau, J., and Rigaud, T. 2000. Operational sex ratio in terrestrial isopods: interaction between potential rate of reproduction and *Wolbachia* induced sex ratio distortion. Oikos **91**(3): 477–484. doi:10.1034/j.1600-0706.2000.910308.x.

Otronen, M. 1984. The effect of differences in body size on the male territorial system of the fly *Dryomyza anilis*. Anim. Behav. **32**(3): 882–890. doi:10.1016/s0003-3472(84)80166-9.

Parker, G.A., Baker, R.R., and Smith, V.G.F. 1972. The origin and evolution of gamete dimorphism and the male-female phenomenon. J. Theor. Biol. **36**(3): 529–553.

Perlman, S.J., Hodson, C.N., Hamilton, P.T., Opit, G.P., and Gowen, B.E. 2015. Maternal transmission, sex ratio distortion, and mitochondria. PNAS **112**(33): 10162–10168. doi:10.1073/pnas.1421391112.

Presgraves, D.C., Severance, E., and Wilkinson, G.S. 1997. Sex Chromosome Meiotic Drive in Stalk-Eyed Flies. Genetics **147**(3): 1169–1180.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available from https://www.r-project.org/.

Rigaud, T., and Moreau, J. 2004. A cost of *Wolbachia*-induced sex reversal and female-biased sex ratios: decrease in female fertility after sperm depletion in a terrestrial isopod. Proc. R. Soc. B Biol. Sci. **271**(1551): 1941–1946. doi:10.1098/rspb.2004.2804.

Ryan, M.J., Dries, L.A., Batra, P., and Hillis, D.M. 1996. Male mate preferences in a gynogenetic species complex of Amazon mollies. Anim. Behav. **52**(6): 1225–1236. doi:10.1006/anbe.1996.0270.

Schlupp, I. 1991. Male Mate Choice in Mixed Bisexual / Unisexual Breeding Complexes of Poecilia (Teleostei : Poeciliidae). **222**: 215–222.

Schlupp, I., and Plath, M. 2005. Male mate choice and sperm allocation in a sexual/asexual mating complex of Poecilia (Poeciliidae, Teleostei). Biol. Lett. **1**(2): 169–171. doi:10.1098/rsbl.2005.0306.

Schultz, R.J. 1969. Hybridization, Unisexuality, and Polyploidy in the Teleost Poeciliopsis (Poeciliidae) and Other Vertebrates. Am. Nat. **103**(934): 605–619.

Stouthamer, R., Breeuwer, J.A.J., Luck, R.F., and Werren, J.H. 1993. Molecular identification of microorganisms accociated with parthenogenesis. Nature **361**: 66–68.

Sullivan, J., and Jaenike, J. 2006. Male-killing *Wolbachia* and male mate choice: A test with *Drosophila innubila*. Evol. Ecol. Res. **8**(1): 91–102.

Thornhill, R. 1976. Sexual Selection and Paternal Investment in Insects. Am. Nat. **110**(971): 153–163.

Trivers, R.L. 1972. Parental investment and sexual selection. *In* In: Sexual Selection and the Descent of Man (ed. Campbell B), London: Heinemann. pp. 136–179. doi:10.1002/ajpa.1330400226.

Uzzell, T.M. 1964. Relations of the Diploid and Triploid Species of the Ambystoma jeffersonianum Complex (Amphibia. Copeia **1964**(2): 257–300.

Vrijenhoek, R.C. 1994. Unisexual fish: model systems for studying ecology and evolution. Annu. Rev. Ecol. Syst **25**: 71–96.

Wedell, N., and Arak, A. 1989. The wartbiter spermatophore and its effect on female reproductive output (Orthoptera: Tettigoniidae, Decticus verrucivorus). Behav. Ecol. Sociobiol. **24**(2): 117–125. doi:10.1007/BF00299643.

Wedell, N., Gage, M.J.G., and Parker, G.A. 2002. Sperm competition, male prudence and sperm-limited females. Trends Ecol. Evol. **17**(7): 313–320.

Werren, J.H., Skinner, S.W., and Huger, A.M. 1986. Male-killing bacteria in a parasitic wasp. Science. **231**(4741): 990–992.

Wilkinson, G.S., Johns, P.M., Kelleher, E.S., Muscedere, M.L., and Lorsong, A. 2006. Fitness effects of X chromosome drive in the stalk-eyed fly, Cyrt*odiopsis dalmanni*. J. Evol. Biol. **19**(6): 1851–1860. doi:10.1111/j.1420-9101.2006.01169.x.

Wilkinson, G.S., Kahler, H., and Baker, R.H. 1998. Evolution of female mating preferences in stalk-eyed flies. Behav. Ecol. **9**(5): 525–533.

Wilkinson, G.S., and Reillo, R. 1994. Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly. Proc. R. Soc. **255**(1342): 1–6.

Woodhead, A.D., and Armstrong, N. 1985. Aspects of the mating behaviour of male mollies (Poecilia spp.). J. Fish Biol. **27**(5): 593–601.

Chapter 4 – Hypotheses and prospective work

Liposcelis sp. is an exciting system to study the evolution of sex determination methods and sex ratio distortions caused by a selfish genetic element. As this work was an initial investigation into the candidate gene Odile, a number of unanswered questions still remain. Although Odile is a promising candidate for the sex ratio distorter, identifying the gene function could provide information required to make more specific predictions on the mechanism of distortion. It was proposed by Hamilton et al. (2018) that distorter *Liposcelis* sp. are acting as feminized males, as they transmit only their maternal genes to the next generation. As paternal genome elimination is the baseline sex determination method in this species (Hodson et al. 2017), it is predicted that the distorting element has hijacked this system (Hamilton et al. 2018).

Paternal genome elimination has evolved multiple times in arthropods, although the underlying mechanisms in how it is achieved are still unclear. Sanchez (2008) hypothesized that in species that determine sex through PGE, a maternal factor is deposited into an egg destined to develop as a male. The proposed maternal factor would result in the heterochromatinization of paternal chromosomes, while the maternal chromosomes, likely imprinted, avoid elimination. As a consequence, males will develop as functionally haploid and pass on only the maternal genes to the next generation. I hypothesize that the distorting element in distorter *Liposcelis* sp. may prevent the heterochromatinization of the paternal chromosomes in the somatic cells, and feminize males by either inhibiting the maternal factor itself, or a target on the paternal chromosome. For example, if the maternal factor is similar to heterochromatin protein 1 (HP1), then this could recognize and bind preferentially to the paternal chromosomes initiating heterochromatinization. LeRoy et al. (2009) found that phosphorylation of histone H1 prevents HP1 from binding to the nucleosome, thus preventing heterochromatinization. Perhaps the distorting element is acting similarly, phosphorylating histone H1 in distorters and turning the “phospho-switch”. The life-stage transcript levels of Odile in Chapter 2 (Figure 2-8) showed increased transcript abundance in nymphs, before individuals were morphologically female. This supports the

theory that Odile could be acting to prevent heterochromatinization of the paternal gene set in early development. Additionally, as distorters do not pass on paternal genes to their offspring, I hypothesize that the distorting element is expressed somatically but is not expressed in the germline. In this case, if the maternal factor is not inhibited, the paternal genes would be heterochromatinized, and consequently not passed on.

In order to investigate further whether Odile is the distorting element in *Liposcelis* sp., functional analysis such as interference RNA (RNAi) or expression in yeast could be utilized. As the life-stage transcript abundance of Odile presents a pattern of increased abundance in young nymphs, injection of dsRNA may not be feasible. Introducing the dsRNA to distorter nymphs via expression in *E. coli* may be a more reasonable approach. Theoretically, if the knockdown of the Odile gene was successful, we may then expect to see both males and females produced by the distorters. Additionally, expression of Odile in yeast could provide information on subcellular localization or protein-protein interactions if the gene product acts on conserved eukaryotic cellular processes. If Odile is successfully expressed in yeast, co-immunoprecipitation followed by mass spectrometry could be used to identify target proteins providing further information on putative function.

It should be noted that the sampling of this species is limited, as *Liposcelis* sp. was collected from one population in the Chiricahua Mountains in Arizona. In general, little is known about wild *Liposcelis* booklice and it would be beneficial to carry out additional field studies to characterize and examine their evolutionary history further. It would be interesting to look for the presence of the distorting element in other populations of booklice and to compare the evolutionary relationships. Additionally, this survey could include an investigation into the mitochondria of *Liposcelis* as they are known to have multiple, circular mitochondrial chromosomes that are rapidly evolving (Cameron, 2014). Perlman et al. (2015) found that the distorter and the nondistorter *Liposcelis* sp. have five and seven mitochondrial minichromosomes respectively. As both the mitochondria and the distorting element in the distorter *Liposcelis* sp. are maternally transmitted, the two are in perfect linkage with each other resulting in the divergence of the mitochondria from the nondistorter population. As the sequencing of the mitochondrial haplotypes in *Liposcelis* sp. was done using lab populations, screening of

wild populations of *Liposcelis* sp. and other booklouse species for mitochondrial divergence could be very interesting and informative.

Finally, because there is limited information on *Wolbachia* infection status in *Liposcelis*, screening wild booklice for *Wolbachia* could help to elucidate the relationship between Odile and its *Wolbachia* ancestor. *Wolbachia* infection in booklice has thus far only been reported in *Liposcelis tricolor* (Dong et al. 2006), and there is currently no sequence data available for analysis. It would be very interesting to look for an Odile homolog in the *Wolbachia* that infect *L. tricolor,* or any other *Wolbachia* bacterium that is found to infect other booklouse species.

Chapter 4 – References

Cameron, S.L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annual Review of Entomology 59:95–117.

Dong, P., Wang, J., and Zhao, Z. 2006. Infection by *Wolbachia* bacteria and its influence on the reproduction of the stored-product psocid, *Liposcelis tricolor*. J Insect Sci. **6**(1): doi.org/10.1673/2006_06_24.1

Hamilton, P.T., Hodson, C.N., Curtis, C.I., and Perlman, S.J. 2018. Genetics and genomics of an unusual sex ratio distortion in an insect. Curr. Biol. **28**(23): 3864-3870.

Hodson, C.N., Hamilton, P.T., Dilworth, D., Nelson, C.J., Curtis, C.I., and Perlman, S.J. 2017. Paternal Genome Elimination in *Liposcelis* Booklice. Genetics **206**(June): 1091-1100. doi:10.1534/genetics.117.199786/-/DC1.1

LeRoy, G., Weston, J.T., Zee, B.M., Young, N. L., Plazas-Mayorca, M.D., and Garcia, B.A. 2009. Heterochromatin protein 1 is extensively decorated with histone code-like post translational modifications. MCP. **8**(11): 2432-2442.

Perlman, S.J., Hodson, C.N., Hamilton, P.T., Opit, G.P., and Gowen, B.E. 2015. Maternal transmission, sex ratio distortion, and mitochondria. PNAS. **112**(33): 10162–10168. doi:10.1073/pnas.1421391112.

Sanchez, L. 2008. Sex-determining mechanisms in insects. Int. J. Dev. Biol. **52**(7): 837-856.