Chapter 5

The *transformer* gene is duplicated in *Leptopilina clavipes* and regulated differently in arrhenotokous and *Wolbachia*-induced thelytokous individuals

Elzemiek Geuverink
Louis van de Zande
Ken Kraaijeveld
Leo W. Beukeboom
ABSTRACT

A transformer (tra) ortholog (Lc-tra) and a tra paralog (Lc-traB) were identified in the genome of Leptopilina clavipes (Hymenoptera: Cynipidae). Compared to Lc-tra, Lc-traB lacks the region coding for male-specific exons and is not sex-specifically spliced. Two types of reproduction occur in L. clavipes, arrhenotoky and Wolbachia-induced thelytoky. Wolbachia-infected females do not produce male offspring; instead females develop from unfertilized diploid eggs. Thelytokous females provide female-specific Lc-tra to their eggs, but arrhenotokous females do not. In contrast, Lc-traB is maternally provided to the embryo in both arrhenotokous and thelytokous L. clavipes. Regulation of Lc-tra splicing differs however between the reproductive modes. In arrhenotokous adult females, both the female and male splice form of Lc-tra are present in comparable amounts, whereas female-specific splice forms predominate in thelytokous females. Thelytokous females that are cured of Wolbachia infection produce male offspring with male-specific Lc-tra splicing. Apparently, in unfertilized thelytokous eggs female-specific splicing of Lc-tra is effectuated by Wolbachia manipulation of the female developmental pathway.
INTRODUCTION

All Hymenoptera are haplodiploid and do not possess heteromorphic sex chromosomes. Their mode of sexual reproduction is arrhenotoky, i.e. haploid males develop parthenogenetically from unfertilized eggs and diploid females from fertilized eggs. A large number of species can produce diploid females from unfertilized eggs parthenogenetically. This reproduction mode is called thelytoky. Thelytoky can be caused by infection with endosymbionts, but also by nuclear factors (Stouthamer, 1997; Lattorff et al., 2005; Sandrock & Vorburger, 2011). In the cynipid wasp *L. clavipes* both arrhenotokous and *Wolbachia*-infected thelytokous populations exist, which are geographically separated (Pannebakker et al., 2004c; Kraaijeveld et al., 2011). The cytological mechanism of thelytokous reproduction is gamete duplication, i.e. diploidy of the egg is restored by failure of the first mitotic anaphase division (Pannebakker et al., 2004b). This mitotic aberration results in identical chromosome pairs and thus complete homozygosity. The fact that such homozygous diploids develop as females if *Wolbachia* is present, argues against a mechanism of sex determination that involves complementarity of sex alleles (complementary sex determination, CSD (Whiting, 1939, 1943; Cook, 1993)). Under CSD, homo- or hemizygosity at one or more sex loci results in male development. Absence of single locus CSD (sl-CSD) is consistent with no increased sex ratio or male mortality in arrhenotokous strains under inbreeding conditions in the lab (personal observations K. Kraaijeveld). Thelytoky, resulting in homozygous female individuals, would however be possible if the arrhenotokous CSD signal of sex determination can be subverted and become obsolete.

The manipulation of host reproduction by endosymbiotic bacteria, such as *Wolbachia* and *Cardinium*, is widespread in arthropods (O’Neill et al., 1997). These endosymbionts, of which *Wolbachia* is the most common with a presence in 40% of all terrestrial arthropods (Zug & Hammerstein, 2012), can cause cytoplasmic (egg-sperm) incompatibility, male progeny killing, feminization of genetic males and parthenogenetic (all-female) reproduction. The bacteria live intracellularly and are vertically transmitted through egg cytoplasm, but not through sperm. Therefore, bacterial traits that favour female production of their host are under positive selection. Parthenogenesis-inducing endosymbionts achieve the maximal induction of female bias by removing the need for fertilization. Infected female hosts produce only daughters that in turn all transmit the endosymbionts to their daughters (Werren et al., 2008).

The mode of reproduction and mode of sex determination in Hymenoptera are tightly connected. Endosymbionts can manipulate ploidy level through a variety of mechanisms and can also cause feminization separate from a ploidy change (Stouthamer, 1997; Giorgini et al., 2009). Hypothetically diploidization of the haploid egg may be sufficient for the endosymbiont to ensure female development, if femaleness is a direct consequence of diploidy. However, for the parasitoid *Asobara japonica* (Ma et al., 2015) reported evidence for a two-step mechanism consisting of diploidization and feminization. The diploidization itself did not start female development and a higher quantity of *Wolbachia* was needed to overrule the sex determination mechanism, i.e. change diploid males into females. Whether a change from haploidy to diploidy
ensures female development depends on how the ploidy level is processed in the sex determination mechanism. It is not known whether a two-step mechanism of diploidization and feminization is also present in *L. clavipes*.

Despite a fast evolving genetic sex determination cascade, most insect species have functionally conserved the gene *transformer (tra)* as the transducer of the primary signal. Nevertheless, *tra* displays high interspecific sequence divergence (Verhulst *et al.*, 2010b). *Tra* is considered to be ancestral to the holometabolous insects based on the presence of specific sequences that are conserved among insect orders (Verhulst *et al.*, 2010b; Geuverink & Beukeboom, 2014). In the order of the Hymenoptera, *tra* has been found in nearly all families (except in *Athalia rosae* representing the basal lineage of Tenthredoidea (i5K *Athalia* genome project, unpublished data)). In addition, paralogs have been identified in multiple branches of the Aculeata (Schmieder *et al.*, 2012; Privman *et al.*, 2013; Koch *et al.*, 2014). A sex-determining function of a *tra* paralog has thus far only been documented in *Apis mellifera*, where it has been identified as the complementary sex determining locus (*csd*) (Beye *et al.*, 2003; Hasselmann *et al.*, 2008a). *Tra* paralogs have been found predominantly in families that have tested positive for CSD, but a recent study has demonstrated their presence in a suggested non-CSD clade (Jia *et al.*, 2016).

In the present study we use the polymorphic reproductive system of *L. clavipes* (Pannebakker *et al.*, 2004c; Kraaijeveld *et al.*, 2011) to elucidate the sex determination mechanism of arrhenotokous and thelytokous wasps. We report the identification of *tra* homologs from the sequenced genome of *L. clavipes*. Expression patterns of these homologs are analyzed in arrhenotokous and thelytokous eggs and adults. Next, expression is monitored in thelytokous individuals that are cured from *Wolbachia* infection by antibiotic treatment. This will inform us on the manipulation performed by *Wolbachia* to induce thelytoky and provide clues at which level of the sex determination cascade this occurs.

**MATERIAL AND METHODS**

**Identification of *tra* homologs**

Scaffolds containing putative *tra* homologs were identified from the genomic assembly of strain GBW (Kraaijeveld *et al.*, 2016) using the protein sequence of *N. vitripennis tra* (NP_001128299) as a query in translated BLAST (tblastn) (Altschul *et al.*, 1997). Adult males and females of the arrhenotokous strain (EPG) and adult females of the thelytokous strain (GBW) were collected from laboratory cultures. The EPG strain was collected in Calzone (Spain) in 2005 (Kraaijeveld *et al.*, 2011) and the GBW strain originated from Wolfheze (The Netherlands) in 2000 (Pannebakker *et al.*, 2004c). RNA extractions were performed according to manufacturer’s protocol with TriZol (Invitrogen, Carlsbad, California, USA). All isolated total RNA was primed with oligo(dT) and random hexamers (in a mixture of 1:6) and reverse-transcribed with the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas, Hanover, MD, USA). Reverse transcription for 3’RACE adapter synthesis was performed with the RevertAid™ H Minus First
The *tra* gene is duplicated in *L. clavipes* and regulated differently in arrhenotokous and thelytokous individuals.

Strand cDNA Synthesis Kit (Fermentas, Hanover, MD, USA) using all isolated total RNA primed with 3’RACE adapter (5’-GCG AGC ACA GAA TTA ATA CGA CTC ACT ATA GGT 12VN-3’). 5’RACE adapter containing cDNA was produced according to manufacturer’s instructions (FirstChoice RLM-RACE kit, Ambion, Austin, TX, USA). To assess the *Lc-tra* splice variants present in adult males and females 5’RACE-PCR was performed with outer primer Lcla_tra_5RACE1 (5’-ATTGACAAGAAGAAGAAGC-3’) and inner primer Lcla_tra_5RACE2 (5’-CCAGATATGTTTCGTGAAT-3’) in a reaction at 94°C for 3 minutes, 40 cycles of 94°C for 30 seconds, 54°C for 30 seconds and 72°C for 60 seconds, with a final extension of 10 minutes at 72°C. Outer primer Lcla_tra_3RACE1 (5’-TGAAACCTTTGTTCGTGGAC-3’) and inner primer Lcla_tra_3RACE2 (5’-AACATATCTGACCCGTCA-3’) were used in 3’RACE-PCR in a reaction with DreamTaq (Fermentas, Hanover, MD, USA). Cycling conditions were at 94°C for 3 minutes, 35 cycles of 94°C for 30 seconds, 59°C for 30 seconds and 72°C for 2 minutes, with a final extension of 7 minutes at 72°C. Resulting PCR fragments were run and visualized on ethidiumbromide-containing 1.5% agarose gel.

All RACE-PCR products were ligated into pGEM-T vector (Promega, Madison, WI, USA) after purification using GeneJET Gel Purification Kit (Fermentas, Hanover, MD, USA). Ligation reactions were used to transform competent JM-109 *Escherichia coli* (Promega, Madison, WI, USA). Colony-PCR was conducted by use of pGEM-T primers (5’-GTAAACGACGGCCAGT-3’) and (5’-GGAACACGCTATGACCATG-3’) at 94°C for 3 minutes, 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 2 minutes, with a final extension of 7 minutes at 72°C.

The lowered specificity of RACE-PCRs (one gene-specific primer per PCR) rarely permitted the detection of *Lc-traB* thus Reverse Transcription (RT-)PCRs were used to detect splice variation in this gene. These PCRs were performed with primers Lcla_traB_frontF (5’-GAGACAAGAAGAAGAAGC-3’) and Lcla_traB_endR (5’-TGATTCTGTCCTAAGGTGA-3’) in a reaction at 94°C for 3 minutes, 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 2 minutes, with a final extension of 7 minutes at 72°C.

PCR-fragments were sequenced on an ABI 3730XL (Applied Biosystems) and reads were inspected in Chromas (Technelysium) and aligned in MEGA4 (Tamura et al., 2007). Exon-intron structure of the genes was constructed by comparing the mRNA sequences to the PacBio assembly scaffolds (*Lc-tra*: scf7180005166757, *Lc-traB*: scf7180005164248) and visualized with Exon-Intron Graphic Maker (http://wormweb.org/exonintron).

**Phylogeny and haplotype construction**

The following sequences were used in alignments and phylogenetic analyses: *Apis mellifera fem* (AAS86667) and *csd* (AAS86653), *Apis dorsata fem* (ABV56232) and *csd* (ABV36165), *Apis cerana fem* (ABV56230) and *csd* (ABV58877), *Bombus terrestris traA* (NP_001267853) and *traB* (XP_003394693), *Bombus impatiens traA* (XP_003493796) and *traB* (XP_003491525), *Nasonia vitripennis tra* (NP_001128299). Ant protein sequences were obtained from (Privman et al., 2013) supplementary materials. The *L. clavipes* TRA amino acid sequence was blasted.
(translated blast, (Altschul et al., 1997)) against transcriptome shotgun assemblies of *Leptopilina heterotoma* and *Leptopilina boulardi* (Goecks et al., 2013). Fragments of *L. heterotoma* tra (GAJC01007654) and traB (GAJC01007653 and GAJC01007652) and *L. boulardi* tra (GAJA01009508) were obtained from PRJNA202370 and PRJNA202369. A neighbour-joining tree was constructed in MEGA4 (Tamura et al., 2007) and alignments were produced in CLC workbench (CLCbio, Aarhus, Denmark).

DNA extractions were performed with a standard high salt protocol (Aljanabi & Martinez, 1997). To assess variation in the *tra* genes between different populations of *L. clavipes* 5 females per strain were individually used for DNA extraction obtained from frozen samples of 12 thelytokous strains (AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, KBH, MGS, STP, WB1a, WB3) and 9 arrhenotokous strains (CBY, DC, EJ, EPG, Mol, MS, PdA, PLB, TL) as described in (Pannebakker et al., 2004c; Kraaijeveld et al., 2011). Population variation was tested with primer Lcla_traA_F (5’-GTCCATCATTCAGAGACAGAC-3’) in combination with Lcla_traA_R (5’-AGGTCATTATTTATCGACGG-3’) and Lcla_traB_R (5’-AGGT CATTATTTACAATGATGG-3’). Reaction conditions were 94°C for 3 minutes, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 45 seconds, with a final extension of 7 minutes at 72°C. Fragments were sequenced, inspected in Chromas (Technelysium) and aligned in MEGA4 (Tamura et al., 2007). A median-joining haplotype network was constructed with PopART (http://popart.otago.ac.nz).

**Lc-tra splicing in arrhenotokous and thelytokous individuals**

Adults of an arrhenotokous strain (ATL) and adult females of a thelytokous strain (KBH) were collected from lab cultures. The ATL strain was collected in 2011 in Atlanta (USA) (http://people.reed.edu/~schlenkt/resources.html) and the KBH strain in ‘s-Heerenberg (The Netherlands) in 2000 (Pannebakker et al., 2004c). The subsequent experiments were performed on these strains as all other *L. clavipes* lab cultures had been discontinued prior to this project. Females of the thelytokous strain KBH were cured from their *Wolbachia* infection by feeding honey with 0.5% rifampicin; this results in haploid eggs that develop into males (referred to as “thelytokous males”, (Schidlo et al., 2002)). Females were allowed to parasitize second instar *Drosophila subobscura* larvae for a period of 4 hours. 10 parasitized *Drosophila* larvae containing *L. clavipes* embryos were collected per sample (4 samples for thelytokous and 4 samples for mated arrhenotokous wasps) for use in RNA extractions. RNA extractions of embryos (parasitized *Drosophila*) and adults were performed according to manufacturer’s protocol with TriZol (Invitrogen, Carlsbad, California, USA). All isolated total RNA was primed with oligo(dT) and random hexamers (in a mixture of 1:6) and reverse-transcribed with the RevertAidTM H Minus First Strand cDNA Synthesis Kit (Fermentas, Hanover, MD, USA).

The presence of sex-specific splice variants of *Lc-tra* in arrhenotokous males and females and thelytokous males and females was tested with primers Lcla_tra_spliceA_F (5’- CAGTCAGAGACAGAGATCC-3’) and Lcla_tra_spliceA_R (5’-TACTTCTGATGTTCACTTCC-3’) TACTTCTGATGTTACCTTCC. Presence of *Lc-traB* was detected with primers Lcla_traspecificF
The *tra* gene is duplicated in *L. clavipes* and regulated differently in arrhenotokous and thelytokous individuals.

(5′-CAAGCAGAAGCAAAGATAGAACCA-3′) and Lcla_traspecificR (5′-TTACTTGCCTTGACTTCA-3′). The cycling-conditions were 94°C for 3 minutes, 45 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 2 minutes, with a final extension of 7 minutes at 72°C. Resulting fragments of each category were sequenced to verify their identity as *Lc-tra* male- and female-specific splice variants and *Lc-traB*. Amplification of *Lc-traM* transcripts in adults was performed with primers Lcla_tra_exon3F (5′-TGAACCTTTGTCGTCGGA-3′) and Lcla_tra_MR (5′-TCTAGTTTCCTCATCTTTC-3′) at 94°C for 3 minutes, 40 cycles of 94°C for 30 seconds, 54°C for 30 seconds and 72°C for 30 seconds, with a final extension of 7 minutes at 72°C.

RESULTS

Structure of *tra* in *L. clavipes*

Two homologous sequences of *transformer* were found in *L. clavipes*, located on two separate genomic scaffolds. The two homologs are present in both arrhenotokous and thelytokous strains. Although the loci have distinctly different genomic structure (Figure 5.1), they code for similar mRNA sequences. *Lc-tra* has sex-specific splice variants, which match those of known *transformer* genes with a function in sex determination. The female splice variant of *Lc-tra* (*Lc-traF*) contains all known functional domains of hymenopteran *transformers*: the Hymenoptera domain, the CAM domain, an arginine/serine-rich region and a proline-rich region (Figure 5.2). The predominant male-specific splice variant (*Lc-traM1*) results in a premature STOP-codon shortly after the Hymenoptera domain. Another, less abundant male specific splice variant *Lc-traM2* contains a STOP-codon at the same position, but merges the sixth and seventh exon (Figure 5.2).

In contrast to *Lc-tra*, *Lc-traB* is not differentially spliced and the locus lacks the male specific exon sequences and some intronic sequences. The single splice variant of *Lc-traB* contains an ORF which closely matches the female-specific splice variant of *Lc-tra*. The conserved domains

![Figure 5.1](image-url)
Figure 5.2. Exon-intron structure of the female-specific and male-specific splice variants of *Lc-tra* and the non-specific splice variant of *Lc-traB*. White boxes represent the 5’ and 3’UTR’s, the black boxes contain the coding sequence. Grey boxes depict the Hymenoptera specific domain (HYM) and the putative autoregulatory region (CAM). The alternative splice form of *Lc-traM* is indicated by the vertical dotted lines connecting the sixth and seventh exon. The scale bars at the right represent 100bp.

of *transformer* are present in the *Lc-traB* homolog (Hymenoptera domain and CAM domain plotted in figure 5.2). This transcript is present in both males and females. The most conserved domain of *transformer*, a putative autoregulatory region referred to as CAM-domain, displays strong divergence in *Lc-traB* (Figure 5.3). Based on these results, we conclude that *Lc-tra* is the *Leptopilina* *transformer* ortholog and *Lc-traB* a paralog of *Lc-tra*.

Conservation of *tra* in the *Leptopilina* genus

The two *tra* homologs of *L. clavipes* are more similar to each other than to any other hymenopteran *tra* homolog (Figure 5.4). This matches a pattern observed in bumblebees and ants (Schmieder et al., 2012; Privman et al., 2013; Koch et al., 2014). The *tra* homolog in the honeybee species (*A. mellifera, A. cerana, A. dorsata*) is called *feminizer (fem)* and is duplicated. This paralog contains a hypervariable region and was identified as the *complementary sex determiner (csd)* locus (Hasselmann et al., 2008b). The hypervariable region is not present in *tra* paralogs of bumblebees and ants, and also does not appear in either of the *L. clavipes tra* duplicates.
The tra gene is duplicated in *L. clavipes* and regulated differently in arrhenotokous and thelytokous individuals.

**Figure 5.3.** Alignment of the CAM domain (putative autoregulatory region) among Hymenoptera. The first 10 amino acids of TRA orthologs are present in both males and females, the latter 15 amino acids are only translated from the female-specific splice variant. CSD and TRAB contain a full CAM domain in both sexes. *Amel* = *Apis mellifera*, *Bter* = *Bombus terrestris*, *Aech* = *Acromyrmex echinatior*, *Acep* = *Atta cephalotus*, *Pbar* = *Pogonomyrmex barbatus*, *Cflo* = *Camponotus floridanus*, *Nvit* = *Nasonia vitripennis*, *Lhet* = *Leptopilina heterotoma*, *Lbou* = *Leptopilina boulardi* and *Lcla* = *Leptopilina clavipes*. TRA of *L. heterotoma* is excluded because the contig only contains the gene up to the CAM domain.

Two *tra* homologs are detected in *L. heterotoma*, while only one fragment was retrieved from the *L. boulardi* dataset, which is similar to the *Lc-tra* CAM domain (Figure 5.3). The missing duplicate in *L. boulardi* is likely an artifact of the dataset, based on transcripts of abdomens only (Goecks et al., 2013). The *tra* sequence of *L. heterotoma* is fragmentary and does not cover the CAM region depicted in Figure 5.4. The available fragments of *tra* and *traB* in *L. heterotoma* cluster together similar to the *L. clavipes* within-species conservation (Figure 5.4).

**Divergence of tra and traB in thelytokous and arrhenotokous L. clavipes**

Fragments of the two *tra* homologs were sequenced for a range of arrhenotokous and thelytokous *L. clavipes* strains, used priorly for genetic diversity assays (Pannebakker et al., 2004c; Kraaijeveld et al., 2011). A coding region upstream of the sex-specifically spliced exons in *Lc-tra*, containing 2 non sex-specific exons separated by an intron, was amplified in both arrhenotokous and thelytokous individuals. 33 SNPs and a 3bp deletion separate the two *tra* copies (Figure 5.5). The strains were screened for intra-strain variation of *Lc-tra* or *Lc-traB*, but none was detected which might in the arrhenotokous strains be due to bottlenecks during the prolonged laboratory culturing. The nucleotide polymorphisms in the two *tra* copies were used to assess the genetic divergence between the strains. As is shown in figure 5.5 *Lc-tra*
polymorphisms between strains are resolved into one cluster of arrhenotokous and one cluster of thelytokous variants, with the exception of KBH (which is also an outlier in (Kraaijeveld et al., 2011)). The arrhenotokous and thelytokous \textit{Lc-tra} haplotype are separated by a single non-
synonymous SNP. The separation in an arrhenotokous and a thelytokous haplotype is not found for \textit{Lc-traB}, where 3 haplotypes are detected (Figure 5.5). The arrhenotokous populations, except EPG, share the same haplotype of \textit{Lc-traB}. The thelytokous strains are divided into two clusters with 2 non-
synonymous and 1 synonymous SNP separating their \textit{Lc-traB} haplotypes. Notably, the \textit{Lc-traB} haplotype also present in one arrhenotokous population (EPG) contains a longer intron, that differs from the length of the \textit{Lc-tra} intron by 1bp only, while the other \textit{Lc-}
The *tra* gene is duplicated in *L. clavipes* and regulated differently in arrhenotokous and thelytokous individuals.

**Figure 5.5.** Median-joining haplotype network of *Lc-tra* and *Lc-traB* in the arrhenotokous and thelytokous populations. Population names are noted next to each cluster. Samples were taken of 12 thelytokous strains (AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, MGS, STP, WB1a, WB3) and 9 arrhenotokous strains (CBY, DC, EJ, EPG, Mol, MS, PdA, PIB, TL) as described in (Pannebakker *et al.*, 2004c; Kraaijeveld *et al.*, 2011). Numbers between brackets show the nucleotide differences between each cluster.

*traB* sequences are 76bp shorter. The separation of thelytokous populations in two clusters is also found with neutral markers (microsatellites) and mtDNA (Kraaijeveld *et al.*, 2011). These results are another indication that, not *Lc-traB*, but *Lc-tra* is the *Leptopilina transformer* ortholog. They also suggest that the different reproductive modes have distinct selection pressures on the *Lc-tra* sequence.

**Differential splicing of *Lc-tra* in arrhenotokous and thelytokous individuals**

In adult arrhenotokous males *Lc-tra* is spliced solely into the male variant, whereas arrhenotokous females contain a mix of the male and the female splice variant (Figure 5.6). *Wolbachia*-infected thelytokous females almost exclusively transcribe the female splice variant of *Lc-tra*. Thelytokous males, whose mothers were cured of their *Wolbachia* infection, but who
do not possess any arrhenotokous genetic material, display a male-only splicing pattern. When a RT-PCR is forced on the male-specific exons, all adult wasps, including thelytokous females, will amplify \(Lc-tra^M\) approximately equally (Figure 5.7). This suggests an elevated expression of \(Lc-tra^F\) in thelytokous females. Alternatively, \(Lc-tra\) mRNA levels are similar between thelytokous and arrhenotokous females, but splicing of \(Lc-tra\) is enhanced towards the female-specific splice variant in thelytokous females.

**Figure 5.6.** Presence of female (\(Lc-tra^F\)) and male (\(Lc-tra^M\)) specific splice variants of \(Lc-tra\) and non-sex-specific splice variants of \(Lc-traB\) in early embryos and adults. Depicted are early embryos of thelytokous wasps (developing into *Wolbachia* carrying females) and early embryos of mated arrhenotokous wasps (developing into females and males without *Wolbachia* infection). The adults used are thelytokous females (*Wolbachia* present), thelytokous males (*Wolbachia* absent), arrhenotokous females (*Wolbachia* absent) and arrhenotokous males (*Wolbachia* absent). The negative control contains no cDNA (NTC).

\(Lc-traB\) mRNA, having only one splice variant, is present in all adult individuals (Figure 5.6). This is also true for early embryos, regardless of reproductive mode, and results from either maternal provision or early zygotic transcription of \(Lc-traB\). In contrast, \(Lc-tra\) regulation depends on the reproductive mode. It is present in the female splice form (\(Lc-tra^F\)) in 0-4h old embryos of *Wolbachia*-infected thelytokous mothers, whereas embryos of arrhenotokous females only possess the male splice form (\(Lc-tra^M\)) (Figure 5.6). Moreover, these (uninfected) arrhenotokous embryos only contain \(Lc-tra^M\) mRNA before 4 hours of development. Yet, control samples confirm that they will develop as females that possess the \(Lc-tra^F\) splice form as adults. This indicates that maternal provision of \(tra^F\), or potentially any type of maternal provision, is not required for female development in arrhenotokous *L. clavipes*.
The tra gene is duplicated in *L. clavipes* and regulated differently in arrhenotokous and thelytokous individuals |

Figure 5.7. Presence of *Lc-tra* mRNA in adult arrhenotokous and thelytokous females and males. The negative control contains no cDNA (NTC).

**DISCUSSION**

*Function of Lc-tra and Lc-traB*

*Lc-tra* is likely a *transformer* ortholog as it possesses a conserved sex determination function based on its sex-specific splicing, sequence conservation within reproductive type and its high conservation of domains and structure. This study has shown that it is not maternally provided to offspring which is in contrast to *tra* functionality in the majority of insects. Male-specific *Lc-tra* mRNA can still be detected in thelytokous females which is consistent with the conservation of the genomic region of the male-specific exons. While *Lc-traB* has lost the region with male-specific exons, the thelytokous *Lc-tra* locus reveals no detectable losses. Knockdown studies are required to further assess the functionality of *Lc-tra* and *Lc-traB* in the arrhenotokous and thelytokous reproductive systems.

The loss of the male-specific genomic region in *Lc-traB* could be a consequence of the lack of sex-specific splicing of this paralog. This lack of sex-specific splicing in turn may result from neofunctionalization. The loss of intronic regions is not seen in *csd* of *A. mellifera*, where in fact the genomic region of *csd* is larger than that of its paralog *fem* (Hasselmann *et al.*, 2008a; Schmieder *et al.*, 2012). Clustering of *Lc-traB* variants resembles the phylogenetic pattern obtained from mtDNA and microsatellite variation of the different populations. This clustering is not restricted by arrhenotokous or thelytoky reproduction. It might suggest that *Lc-traB* is released from selection and subject to neutral drift. However, if *Lc-traB* is in the process of becoming a pseudogene (as present in other species (Koch *et al.*, 2014)), this divergence is expected to be more pronounced across the entire gene and not restricted to the region which normally contains male-specific exons. Additionally, if there was no selection to preserve the coding sequence, each *Leptopilina* species *traB* copy should not cluster so closely to its *tra* paralog in the gene tree. This suggests that *traB* has obtained a new function, but the current
information is insufficient as to which function, and whether this is a function in sex determination. (Hasselmann et al., 2010) demonstrated a stronger purifying selection on fem in presence of csd, suggesting that the functional link between csd and fem leads to restrictions on the evolution of the latter. Csd in its turn possesses an excess of shared non-synonymous mutations (Hasselmann et al., 2008b). Similar analyses can be conducted on the yet to be completed tra and traB sequences in the Leptopilina genus and on populations of different reproductive modes in L. clavipes.

Duplications of tra

Paralogs of tra have been linked to CSD systems while little is known about their functionality (Schmieder et al., 2012; Privman et al., 2013). The duplication of tra appears to be present in each Leptopilina species investigated. Two tra homologs can be detected in a transcriptome assembly of L. heterotoma. L. boulardi only contains one homolog in its assembly, but traces of a duplicate Hymenoptera domain appear to be present in the raw read data. Lack of sequence variation resolution was also plaguing the L. clavipes genome assembly prior to the addition of long read data (Kraaijeveld et al., 2016). It suggests that the gene duplication preceded the emergence of the Leptopilina genus. The clustering by species instead of by paralog (Figure 5.3) does not necessarily indicate repeated duplication events after speciation, instead this could be a case of convergent evolution, as has been proposed for other tra duplications in the Aculeata (Schmieder et al., 2012; Privman et al., 2013). These ideas have however been disputed by (Koch et al., 2014), who suggest that the duplications are all separate events. The Leptopilina species are however much more closely related than the aculeatan species which span multiple families. The tra paralogs were until recently only detected in species with a CSD mode of sex determination. This is potentially reflecting a bias in study effort, rather than a true link to CSD mechanisms, as Aculeata are overrepresented in the available hymenopteran genomes. Recently, three homologs of tra were found in the fig wasp Ceratosolen solmsi, a species belonging to the Chalcidoidea in which CSD appears absent (van Wilgenburg et al., 2006; Heimpel & de Boer, 2008; Jia et al., 2016).

Effect of Wolbachia on Lc-tra expression

Wolbachia needs to act early in the sex determination cascade to ensure female development. The exaggerated female-specific splicing of Lc-tra in thelytokous females suggests that Wolbachia interacts with either Lc-tra itself, or an as-of-yet unidentified factor upstream of Lc-tra in the sex determination cascade. To further elucidate Wolbachia’s interaction with putative sex determination genes, the presence of Lc-traB and splicing of Lc-tra was assessed in early embryos. Interestingly Lc-traB was provided to the embryos, but did not show any differences in the level of maternal provisioning between reproductive modes. This indicates a function during early development which may still include a role in sex determination. Although this study demonstrates that thelytoky is associated with female-specific Lc-tra provision, experiments
The tra gene is duplicated in L. clavipes and regulated differently in arrhenotokous and thelytokous individuals | that remove Wolbachia from thelytokous females need to be performed to test whether it is solely the endosymbiont that causes this change to maternal Lc-tra^F input. A genetic factor in the thelytokous genome itself cannot be excluded based on these data. The thelytoky restricted provision of Lc-tra^F suggests that Wolbachia manipulate tra regulation in thelytokous females. These females display an elevated or biased expression of Lc-tra^F, leading to a change in mRNAs provided to the embryos. To test this suggested link between Wolbachia presence and amount of Lc-tra^F, one could change the titer of Wolbachia in thelytokous females and compare it to the expression of Lc-tra in the same individuals.

Embryos of L. clavipes exhibit delayed development and the first meta-anaphase is completed only after 30 minutes following oviposition, but can also extend to 3.5h after oviposition (Pannebakker et al., 2004b). As more than half of the embryos are still delayed at 2h after oviposition (Pannebakker et al., 2004b) and females often do not oviposit immediately when placed in the test set-up, it is likely that a significant fraction of the 0-4h old embryo samples in our study had not reached the stage of zygotic transcription yet. Maternal RNAi tests can be used to assess whether Wolbachia-induced thelytoky impacts maternal provisioning, rather than early zygotic transcription of Lc-tra. These would knock down expression of Lc-tra in adult females, leading to an inability to provide Lc-tra mRNA to their offspring.

Diploidization and feminization in L. clavipes

In A. japonica Wolbachia induced thelytoky follows a two-step model, consisting of separate processes of diploidization and feminization (Ma et al., 2015). It is unknown whether Wolbachia in L. clavipes deploys a similar mechanism. Ploidy screens on the offspring of antibiotics treated females would be needed to detect the presence of diploid males and a possible correlation between Wolbachia titer and ploidy. If feminization is a separate process it would allow a wider range of possible sex determination mechanisms in L. clavipes, including CSD and imprinting mechanisms that are difficult to reconcile with the gamete duplication which underlies its thelytoky. The homozygosity at a genetic or epigenetic level could then be overruled by the endosymbiont through its feminizing effect. Maternal provisioning of tra^F could be a necessary element for the feminization of the thelytokous embryos. Since Lc-tra^F provision is not required for arrhenotokous female development, suggesting that Wolbachia interferes by circumventing the arrhenotokous primary signal, while exploiting tra autoregulation.

How do endosymbionts manipulate host reproduction?

Evidence is growing that the sex determination cascade can be manipulated by endosymbionts, but very little is still known about the underlying mechanisms (Beukeboom, 2012). In the moth Ostrinia scapulalis, the presence of Wolbachia is associated with female-specific splicing of the downstream master-switch gene doublesex (dsx), indicating that this Wolbachia carries a feminizing factor that overrules upstream sex-determination processes (Sugimoto & Ishikawa, 2012). In this system Wolbachia causes male-killing rather than thelytokous reproduction. In the
related moth *Ostrinia furnacalis*, *Wolbachia* represses its host’s masculinizing gene *Masc* (Fukui et al., 2015), which is the target of piRNA derived from the dominant feminizing gene *Fem* in *Bombyx mori* (Kiuchi et al., 2014). This suggests that *Wolbachia* can interfere with the sex determination cascade of *O. furnacalis*. However, the sex determination cascade of these lepidopterans differs notably from other insects in its absence of *tra*. Maternal *tra* provision is only absent after the primary signal addition of *sex-lethal* in *Drosophila* and the *csd* heterozygous state in *A. mellifera*, whereas other insects retain maternal *tra* provision to start the *tra* loop (Bopp et al., 2014). Maternal *tra* provision is of particular importance in *N. vitripennis* (Verhulst et al., 2010a) which is the most closely related investigated species to *L. clavipes*. In *N. vitripennis*, female-specific *tra* mRNA provided by the mother to her embryos, combined with a paternal activator, is required for female development (Verhulst et al., 2013). *L. clavipes* differs from this system in its lack of female-specific *Lc-tra* provision in arrhenotokous females. This has both implications for the possible mechanism of sex determination in *L. clavipes* and the options for *Wolbachia* to interfere with the mechanism. The lack of *tra* provision in arrhenotokous *L. clavipes* may provide opportunities for *Wolbachia* to jumpstart the female pathway by ensuring *tra* provisioning to thelytokous embryos. A *tra* autoregulatory loop is suggested to be widely conserved, activated by maternal provision of *tra* (Bopp et al., 2014).

In individuals without *tra* maternal provision another factor needs to activate *tra* splicing, but the autoregulatory loop of *tra* could already be present in *L. clavipes* at the post-zygotic transcription stages during arrhenotokous diploid (female) development. If *Wolbachia* induces maternal provision of *tra*, it could activate this otherwise still dormant system at the start of zygotic transcription. We examined the thelytokous wasp *Asobara japonica* in chapter 6 to study the extent of *Wolbachia*-induced maternal provision of *tra* and to formulate a model on arrhenotokous and thelytokous sex determination mechanisms in both species.

**Acknowledgements**

We thank Rogier Houwerzijl and Peter Hes for assistance with wasp and fly culturing and Todd Schlenke for supplying the ATL and KBH strains.