Evidence for involvement of a transformer paralog in sex determination of the wasp Leptopilina clavipes

Geuverink, Elzemiek; Kraaijeveld, Ken; van Leussen, Marloes; Chen, Fangying; Pijpe, Jeroen; Linskens, Maarten H K; Beukeboom, Leo W; van de Zande, Louis

Published in:
Insect Molecular Biology

DOI:
10.1111/imb.12522

Publication date:
2018

Insect Molecular Biology

Insect Molecular Biology (2018) **27**(6), 780–795 doi: 10.1111/imb.12522

Evidence for involvement of a *transformer* paralogue in sex determination of the wasp *Leptopilina clavipes*

E. Geuverink*, K. Kraaijeveld†,‡, M. van Leussen*, F. Chen*, J. Pijpe§, M. H. K. Linskens¶, L. W. Beukeboom* and L. van de Zande*

*Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands; †Department of Ecological Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; ‡Leiden Genome Technology Center, Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; §University of Applied Sciences Leiden, Leiden, The Netherlands; and ¶Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

Abstract

***Transformer* (*tra*) is the central gear in many insect sex determination pathways and transduces a wide range of primary signals. Mediated by *transformer-2 (tra2)* it directs sexual development into the female or male mode. Duplications of *tra* have been detected in numerous Hymenoptera, but a function in sex determination has been confirmed only in *Apis mellifera*. We identified a *tra2* orthologue (*Lc-tra2*), a *tra* orthologue (*Lc-tra*) and a *tra* paralogue (*Lc-traB*) in the genome of *Leptopilina clavipes* (Hymenoptera: Cynipidae). We compared the sequence and structural conservation of these genes between sexual (arrhenotokous) and asexual all-female producing (thelytokous) individuals. *Lc-tra* is sex-specifically spliced in adults consistent with its orthologous function. The male-specific regions of *Lc-tra* are conserved in both reproductive modes. The paralogue *Lc-traB* lacks the genomic region coding for male-specific exons and can only be translated into a full-length TRA-like peptide sequence. Furthermore, unlike LC-TRA, the LC-TRAB interstrain sequence variation is not differentiated into a sexual and an asexual haplotype. The LC-TRAB protein interacts with LC-TRA as well as LC-TRA2. This suggests that *Lc-traB* functions as a conserved element in sex determination of sexual and asexual individuals.**

Keywords: *transformer* orthologue, *transformer-2*, Hymenoptera, protein interactions, reproductive modes

First published online 1 October 2018.

Correspondence: Elzemiek Geuverink, Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, 9700 CC Groningen, Groningen, The Netherlands. e-mail: e.geuverink@rug.nl

Introduction

Sex determination is a ubiquitous developmental process in eukaryotes. It entails the differentiation of two sexual functions and leads to the development of female and male morphologies and behaviours. Being a basic developmental process, sex determination may be expected not to tolerate modifications in the underlying developmental pathway as these would disrupt the correct specification of the two sexes. Sex determination is nevertheless characterized by a wide variety of fast-evolving mechanisms, including duplication and subsequent recruiting of sex-determining genes (Beukeboom and Perrin, 2014; Herpin and Schartl, 2015).

Insects constitute a particularly suitable group for studying the regulation of sex determination as they have shown rapid turnover in sex determination mechanisms. The signal-transducing elements of their sex determination cascade are well conserved, but they exhibit a wide variety of upstream signals (Bopp *et al*., 2014). A hallmark of insect sex determination is sex-specific splicing of the transducing genes *transformer* (*tra*) and *doublesex* (*dsx*). The male splice variants of *tra* include exons with in-frame early STOP-codons, resulting in a truncated TRA protein. The female splice variants code for a TRA protein that belongs to a class of SR-type proteins characterized by regions rich in arginines (R) and serines (S). Despite its conserved function, *tra* displays high sequence divergence amongst insects, possibly as a result of accommodating many different upstream primary signals in the cascade (Verhulst *et al*., 2010b). It contains a number of distinctive domains of which the most conserved is the *Ceratitis-Apis-Musca* (CAM) domain, which is believed to implement the

autoregulatory splicing loop of *tra* (Hediger *et al*., 2010). It has been found in all investigated *tra* orthologues, including those of *Tribolium castaneum*, *Apis mellifera*, *Nasonia vitripennis*, *Asobara tabida* and various dipterans, with the exception of drosophilids (Pane *et al*., 2005; Lagos *et al*., 2007; Ruiz *et al*., 2007; Hasselmann *et al*., 2008a; Hediger *et al*., 2010; Verhulst *et al*., 2010b; Saccone *et al*., 2011; Shukla and Palli, 2012; Geuverink *et al*., 2018). TRA also possesses order-specific domains, one of which is only shared amongst the Diptera (the DIP domain) and another that is only present in the Hymenoptera (the HYM domain) (Verhulst *et al*., 2010b). Despite the rapid evolution of sex determination cascades, most insect species have a functionally conserved *tra* orthologue as the transducer of the primary signal.

TRA and Transformer-2 (TRA2) form a complex that controls the sex-specific splicing of *dsx* pre-messenger RNA (pre-mRNA; Nagoshi *et al*. 1988; Nagoshi and Baker, 1990; Hedley and Maniatis, 1991; Hoshijima *et al*. 1991; Inoue *et al*. 1992; Tian and Maniatis, 1992, 1993 ; Amrein *et al*., 1994). TRA and TRA2 are essential components of the autoregulatory *tra* loop and required for splicing of *dsx* transcripts in *T. castaneum*, various dipterans, *Ap. mellifera* and *N. vitripennis* (Pane *et al*., 2002; Lagos *et al*., 2007; Ruiz *et al*., 2007; Hasselmann *et al*., 2008a; Salvemini *et al*., 2009; Sarno *et al*., 2010; Nissen *et al*., 2012; Shukla and Palli, 2012; Shukla and Palli, 2013; Liu *et al*., 2015; Geuverink *et al*., 2017). *Tra2* can be transcribed either into a single isoform or into multiple isoforms that each code for an RNA binding domain (RBD) flanked on both sides by arginine-serine rich regions (Burghardt *et al*., 2005; Niu *et al*., 2005; Concha and Scott, 2009; Salvemini *et al*., 2009; Sarno *et al*., 2010; Martín *et al*., 2011; Nissen *et al*., 2012; Schetelig *et al*., 2012; Shukla and Palli, 2013; Liu *et al*., 2015).

Tra is considered to be ancestral to the holometabolous insects based on its conserved domains (Verhulst *et al*., 2010b; Geuverink and Beukeboom, 2014). However, *tra* orthologues appear absent in some insect groups, notably the Lepidoptera (Salvemini *et al*., 2013; Geuverink and Beukeboom, 2014). In the order of the Hymenoptera, *tra* has been found in nearly all families [except *Athalia rosae*, which represents the basal lineage of Tenthredoidea (Mine *et al*., 2017)]. In addition, paralogues of *tra* 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* paralogue has thus far only been documented in *Ap. mellifera,* where it constitutes the complementary sex-determining locus (*csd*) (Beye *et al*., 2003; Hasselmann *et al*., 2008a). Interestingly, *tra* paralogues have been found predominantly in families that have tested positive for a complementary sex determination system (CSD), although a recent study has demonstrated *tra* paralogues in the suggested non-CSD clade of Chalcidoidea (Jia *et al*., 2016).

All Hymenoptera reproduce by haplodiploidy. Their sexual mode of reproduction is arrhenotoky, in which haploid males develop from unfertilized eggs and diploid females from fertilized eggs. A large number of hymenopteran species are known to reproduce by thelytoky. Thelytokous females produce diploid females from unfertilized eggs parthenogenetically, without a paternal genome contribution. Thelytoky can be the result of infection with endosymbionts, but can also be caused by nuclear factors (Stouthamer, 1997; Lattorff *et al*., 2005; Sandrock and Vorburger, 2011). The manipulation of host reproduction by endosymbiotic bacteria, such as *Wolbachia* and *Cardinium*, is widespread amongst arthropods, in particular amongst Hymenoptera (O'Neill *et al*., 1997).

In the wasp *Leptopilina clavipes* (Hymenoptera; Cynipidae) both arrhenotokous and *Wolbachia*-infected thelytokous populations exist (Pannebakker *et al*., 2004b; Kraaijeveld *et al*., 2011). The cytological mechanism of thelytokous reproduction is gamete duplication, ie diploidy is restored by skipping the first mitotic anaphase division (Pannebakker *et al*., 2004a). This results in identical chromosome pairs and thus complete homozygosity. In northern Europe, *L. clavipes* populations are fixed for *Wolbachia* infection, meaning that they consist of infected thelytokous females only. Conversely, several southern European populations lack this *Wolbachia* infection, and reproduce sexually. Theory predicts that genes with sexual function will degenerate through accumulation of deleterious mutations under asexual reproduction (Kraaijeveld *et al*., 2016). Potential divergence or decay of the sex determination cascade in thelytokous systems has, however, not yet been studied. Furthermore, how endosymbionts achieve their host manipulation is poorly known and requires more knowledge of hymenopteran sex determination mechanisms. The *L. clavipes* system provides this opportunity because the sequences and regulation of sex determination genes can be directly compared between arrhenotokous and thelytokous individuals.

Here we investigate whether and how *tra* and *tra2* function in the sex determination cascade of arrhenotokous and thelytokous lineages of *L. clavipes*. We also screen for paralogues of both genes in both lineages. Splicing patterns of the *tra* orthologue, paralogue and *tra2* are compared in both reproductive modes. Genes are expected to degenerate if they have no function in a particular reproductive mode. Therefore, if genes are only degenerated in thelytokous wasps this would suggest a loss of a sex determining related function at the onset of asexuality induction. However, if these genes are conserved in both reproductive modes and their proteins interact, it would indicate an active function in sex determination. An interaction between TRA and TRA2 is hypothesized

Figure 1. Conservation of the genomic structure of the *Leptopilina clavipes transformer* orthologue (*Lc-tra*) and paralogue (*Lc-traB*). The black lines show the pairwise alignment of *Lc-tra* and *Lc-traB* with the grey block-arrows representing the exons (in the case of *Lc-tra* this shows the female splice variant) and the black block-arrows representing the male-specific exons of *Lc-tra*. The numbers at the right show the length of the genomic *Lc-tra* regions. The 3' untranslated region of *Lc-traB* at the bottom right corner has not been identified. Two genomic regions are not present in *Lc-traB*. The first consists of the male-specific exons in *Lc-tra* and the second matches a *Lc-tra* region without exons.

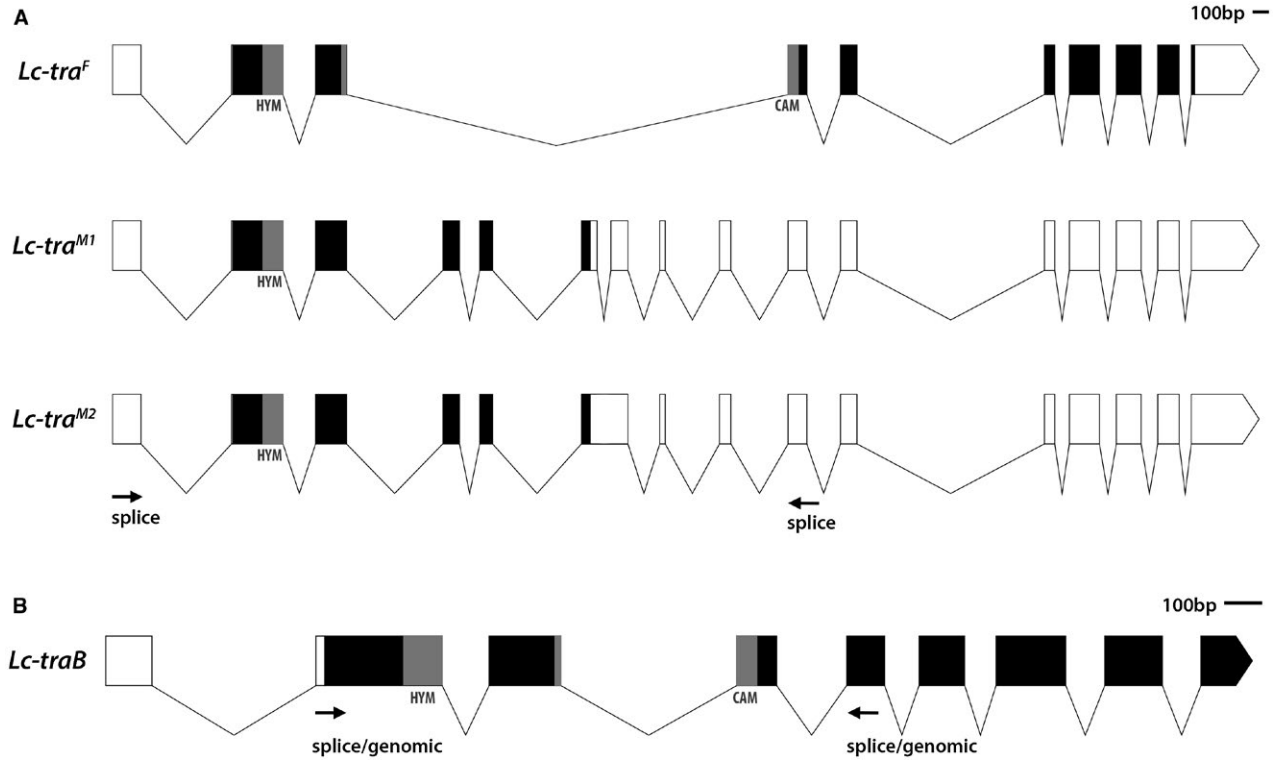

Figure 2. Exon–intron structure of the female-specific and male-specific splice variants of the *Leptopilina clavipes transformer* orthologue (*Lc-tra*) (A) and the nonspecific splice variant of its paralogue *Lc-traB* (B). White boxes represent the 5' and 3' untranslated regions, the black boxes contain the coding sequence. Grey boxes depict the Hymenoptera-specific domain (HYM) and the putative autoregulatory region (*Ceratitis-Apis-Musca*, CAM). The scale bars at the right represent 100 bp. Primer positions for amplification of splicing patterns (Figure 4) and genomic regions (Figure 6) are depicted by arrows underneath each gene structure.

to occur as a requirement for female development. Based on our results, a model for the sex determination system of *L. clavipes* will be presented and compared to known mechanisms within the Hymenoptera.

Results

Identification of tra *homologues and their structure in* L. clavipes

Two homologous sequences of *tra* were found in the *L. clavipes* reference genome assembly [from the thelytokous GBW strain (Kraaijeveld *et al*., 2016)], in two genomic scaffolds (scf7180005166757 and scf7180005164248). The two homologues shared 90.5% identity in their coding region sequence. The two homologues could also be detected in the arrhenotokous wasps by rapid amplification of cDNA ends-PCRs (RACE-PCRs) and reverse-transcription (RT-)PCRs. Although the two loci have a distinctly different genomic structure (Fig. 1), they code for similar mRNA sequences. The gene in scf7180005166757 has sex-specific splice variants that match those of known *tra* genes with a function in sex determination. The female splice variant codes for a peptide of 417 amino acids with all known functional domains of hymenopteran *tras*: the HYM domain, the CAM domain, an arginine-serine (RS)-region and a proline-rich region (Fig. 2). The predominant male-specific splice variant contains a premature STOP-codon shortly after the HYM domain, resulting in a 242-amino-acid protein. Another, less abundant male-specific splice

variant contains a STOP-codon at the same position, but merges the sixth and seventh exons (Fig. 2). Based on these observations we concluded that the gene in scf7180005166757 is the *tra* orthologue and named it *Lc-tra*.

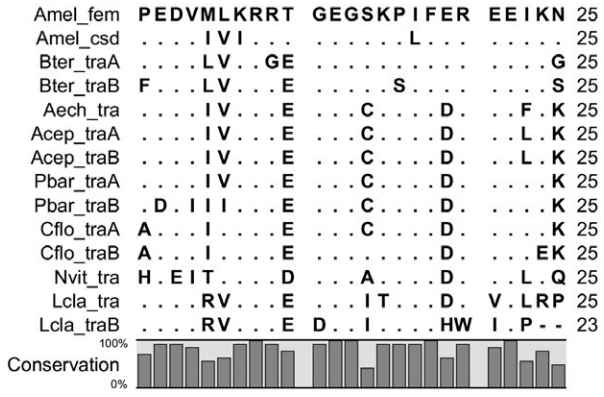

Figure 3. Alignment of the *Ceratitis-Apis-Musca* (CAM) domain (putative autoregulatory region) amongst Hymenoptera. The first 10 amino acids of the transformer (TRA) orthologues are present in both males and females, the latter 15 amino acids are only translated from the female-specific splice variant. CSD and the TRA paralogue (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*; Lcla, *Leptopilina clavipes*; fem, feminizer.

In contrast, the gene in scf7180005164248 is not sex-specifically, indeed not even alternatively, spliced and lacks the region corresponding to male-specific exon sequences and the intron corresponding to that between exons 10 and 11 in *Lc-tra* (Fig. 1). The single splice variant contains an open reading frame (ORF) that closely matches the female-specific splice variant of *Lc-tra*. The conserved domain coding sequences of *tra* are present (HYM domain and CAM domain plotted in Fig. 2), whereas the coding part for the putative autoregulatory region, referred to as the CAM-domain, displays stronger divergence from other hymenopteran sequences (Fig. 3 and Supporting Information Figure S1). Based on these data, we concluded that the gene in scf7180005164248 is a paralogue of *Lc-tra* and named it *Lc-traB*. The peptide sequence measures 429 amino acids and the amino acid sequence similarity is 74% compared to LC-TRAF.

Differential splicing of Lc-tra *and* Lc-traB *in arrhenotokous and thelytokous* L. clavipes

Arrhenotokous *L. clavipes* produce both female and male offspring, whereas thelytokous wasps only produce females. Thelytokous male production can however be induced by antibiotic treatment of the infected females. RNA was extracted from individual adults of each sex and both reproductive types to assess splice variation of sex determination genes. In males of either reproductive

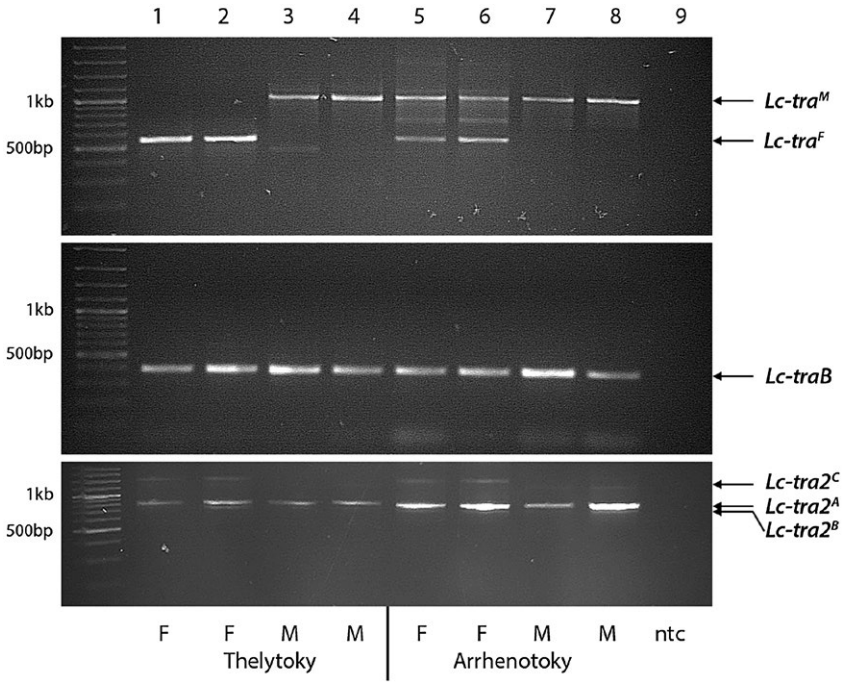

Figure 4. Presence of female (*Lc-tra*^{F}) and male (*Lc-tra*^{M}) specific splice variants of *Leptopilina clavipes transformer* (*Lc-tra*) and non-sex-specific splice variants of its paralogue (*Lc-traB*) and *Lc-tra2* in adult females (F) and males (M) of thelytokous strain KBH (lanes 1–4) and arrhenotokous strain SCA (lane 5-8). The negative control in lane 9 contains no cDNA (ntc).

mode *Lc-tra* pre-mRNA was spliced solely into the male variant, whereas arrhenotokous females contained a mix of the female and the male splice variants (Fig. 4). This is in contrast to the non-treated thelytokous females, which displayed only the female-specific form of *Lc-tra*. The single transcript of *Lc-traB* was abundantly present in females and males of both reproductive modes (Fig. 4). This full-length *Lc-traB* transcript results in the presence of a CAM-like domain coding sequence in males. No splice variation of *Lc-traB* was detected in either sex.

Faint traces of other amplicons were detected in the *Lc-tra* RT-PCR (Fig. 4). The faint band above the *Lc-tra*^{M} arrow was the less abundant *Lc-tra*^{M2} splice variant. The amplicon between *Lc-tra*^{F} and *Lc-tra*^{M} in arrhenotokous individuals and the faint lower band in some thelytokous individuals were not successfully cloned. These potential alternative splice variants could not be detected in the transcriptome of strain EPG, which only contains an isotig of *Lc-tra*^{M1} (accession: GAXY02017594) as well as *Lc-traB* (accession: GAXY02017595) (Peters *et al*., 2017), nor could they be predicted from the genomic sequence.

Sequence divergence of Lc-tra *and* Lc-traB *in thelytokous and arrhenotokous individuals*

Fragments of the two *tra* homologues were sequenced from stored samples of a range of arrhenotokous and thelytokous *L. clavipes* strains, used previously for genetic diversity assays (Pannebakker *et al*., 2004b; Kraaijeveld *et al*., 2011). An overview of these 12 thelytokous strains (AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, KBH, MGS4, STP, WB1a, WB3) and nine arrhenotokous strains (CBY, DC, EJ, EPG, Mol, MS, PdA, PlB, TL) is presented in Supporting Information Table S1.

A coding region upstream of the sex-specifically spliced exons in *Lc-tra*, containing two non-sex-specific exons separated by an intron, was amplified from both arrhenotokous and thelytokous individuals. The two *tra* homologues differ in this region by 33 single nucleotide polymorphisms (SNPs) and a 3-bp deletion (Fig. 5). No intrastrain variation was present in *Lc-tra* and *Lc-traB*. The nucleotide polymorphisms in the two *tra* copies were used to assess the genetic divergence between the lineages. *Lc-tra* polymorphisms between lineages resolve into one cluster of arrhenotokous and one cluster of thelytokous variants, with the exception of lineage KBH (which was also an outlier in Kraaijeveld *et al*., 2011). The arrhenotokous and thelytokous *Lc-tra* haplotypes can be separated by a single nonsynonymous SNP. By contrast, such separation by reproductive mode is not evident for *Lc-traB*, for which three haplotypes were detected (Fig. 5). The arrhenotokous lineages, except EPG, share the same haplotype of *Lc-traB*. The thelytokous lineages are divided into two clusters with two nonsynonymous and one synonymous

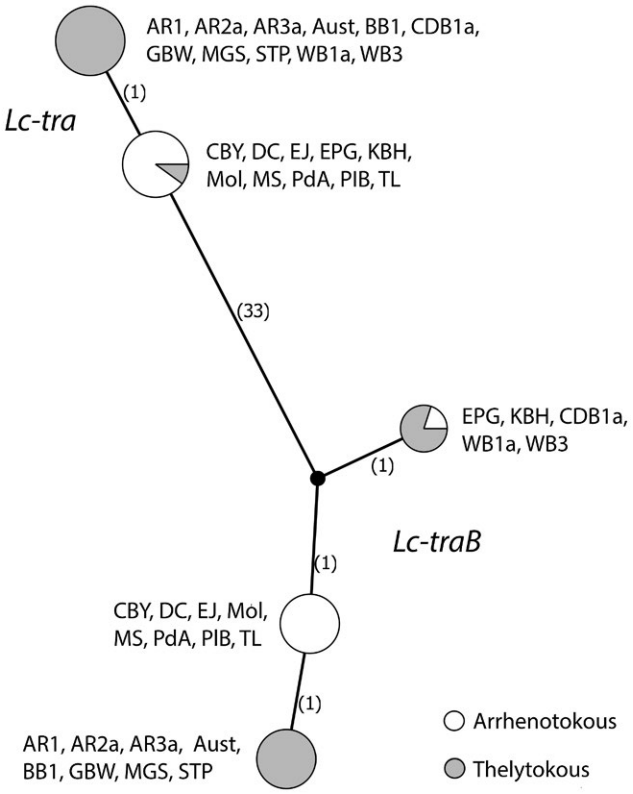

Figure 5. Median-joining haplotype network of *Leptopilina clavipes transformer* (*Lc-tra*) and its paralogue (*Lc-traB*) in the arrhenotokous and thelytokous populations. Population names are noted next to each cluster. Samples were taken of nine arrhenotokous strains (CBY, DC, EJ, EPG, Mol, MS, PdA, PlB, TL) and 12 thelytokous strains (AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, KBH, MGS, STP, WB1a, WB3) as described in Pannebakker *et al*. (2004) and Kraaijeveld et al. (2011). Numbers in parentheses show the nucleotide differences between each cluster.

SNPs separating their *Lc-traB* haplotypes. Notably, the *Lc-traB* haplotype that is found in both arrhenotokous and thelytokous lineages contains a longer intron that is similar in length (1-bp difference) to the *Lc-tra* intron. The other *Lc-traB* haplotypes contain an intron that is 76 bp shorter. The distinction of thelytokous lineages into two clusters was also observed with neutral markers (microsatellites) and mtDNA (Kraaijeveld *et al*., 2011).

The region between exons 3 and 9 of *Lc-tra* contains the male-specific exons that are spliced out in the female form. Thelytokous lineages do not produce males and this region could potentially have degenerated in these lineages without affecting the functionality of *tra*. This large region in *Lc-tra* could not be amplified by PCRs with primers located on exons 3 and 9. As an alternative approach genomic HiSeq data of thelytokous strain MGS4 and arrhenotokous strains EJ and PdA were mapped against the thelytokous GBW reference genome. Comparison of the thelytokous (MGS4/GBW) and arrhenotokous (EJ/PdA) consensus sequences yielded only one intronic

SNP in this male region of *Lc-tra* (Table 1). This indicates that the entire male-specific region is intact in the thelytokous *Lc-tra* lineage (Figs 1, 2). By contrast, the prospective promotor region, the 5· untranslated region (5'UTR), and introns 9 and 10 contain a large number of SNPs and deletions (Table 1). Furthermore, two nonsynonymous SNPs are present between the exonic consensus sequences of the thelytokous and arrhenotokous lineages, but no synonymous SNPs. These patterns confirm the separation of *Lc-tra* into an arrhenotokous and a thelytokous haplotype.

Whereas the thelytokous and arrhenotokous consensus sequences of *Lc-tra* contain a large number of intronic SNPs, *Lc-traB* is almost identical between the two reproductive modes (Table 1). There is some sequence variation amongst strains, but only a single nonsynonymous mutation on exon 2 separates the sequences by reproductive mode. The *Lc-traB* genomic region in the thelytokous *L. clavipes* genome lacks the region that in *Lc-tra* contains male-specific exons (Fig. 1). To examine potential divergence of these regions between the *tra* copies of different lineages, we amplified this intronic region of *Lc-traB* in DNA samples of the arrhenotokous and thelytokous strains. The length of the intronic regions appear conserved in all lineages, regardless of reproductive mode (Fig. 6).

Conservation of Lc-tra *and* Lc-traB

The two *tra* copies of *L. clavipes* are more similar to each other than to any other hymenopteran *tra* homologue (Fig. 7). This matches a pattern observed in bumblebees and ants (Schmieder *et al*., 2012; Privman

Table 1. Comparison of promotor-region, 5' untranslated region (5'UTR), coding DNA sequences (CDS), male-specific region and 3' region between *Leptopilina clavipes transformer* (*Lc-tra*) of arrhenotokous strains (EJ/PdA) and *Lc-tra* of thelytokous strains (GBW/MGS4), the *Lc-tra* paralogue (*Lc-traB*) of arrhenotokous strains and *Lc-traB* of thelytokous strains, and *Lc-tra2* of arrhenotokous and *Lc-tra2* of thelytokous strains. The number of segregating sites includes both single nucleotide polymorphisms (SNPs) and deletions. SNPs located on exons are marked as nonsynonymous or synonymous in the right-most two columns

		Exon+intron #segregating sites	Exon #nonsynonymous mutations	Exon #synonymous mutations
Lc-tra thelytokous vs. arrhenotokous	Promotor region (1000 bp prior to transcription start)	11	N/A	N/A
	5'UTR	14	N/A	N/A
	CDS exons 1–3	1	1 (P in thelytokous vs. S in arrhenotokous)	0
	Male-specific exon region (intron after exon 3 until start exon 9)	1 (intron)	N/A	N/A
	3' (exon 9 until transcript stop)	15 (14 of which in introns)	1 (P in thelytokous vs. A in arrhenotokous)	0
Lc-traB thelytokous vs. arrhenotokous	Promotor region (1000 bp prior to transcription start)	1	N/A	N/A
	5'UTR	0	N/A	N/A
	CDS exons 1–3*	1	1 (K in thelytokous vs. E in arrhenotokous)	0
	Intron between exons 3 and 4	0	N/A	N/A
	3' (exon 4 until transcript stop)	0	0	0
Lc-tra2 thelytokous vs. arrhenotokous	Promotor region (1000 bp prior to transcription start)	4	N/A	N/A
	5'UTR	0	N/A	N/A
	CDS	6 (all intronic)	N/A	N/A
	3'UTR	2 (1 on exon, 1 in intron)	N/A	N/A

interstrain variation in this region, see Figure 4

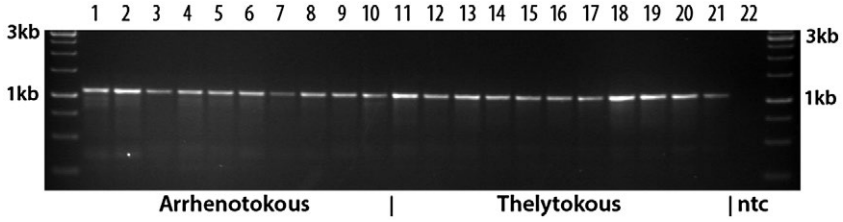

Figure 6. Amplification of intronic regions of the *Leptopilina clavipes* transformer paralogue (*Lc-traB*) in nine arrhenotokous strains (lanes 1–9: CBY, DC, EJ, EPG, Mol, MS, PdA, PIB, TL) and 12 thelytokous strains (lanes 10–21: AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, KBH, MGS, STP, WB1a, WB3). The negative control in lane 22 contains no cDNA (ntc). The amplified fragment is 1129 bp and includes the truncated intronic region in *Lc-traB* that is homologous to the region containing male-specific exons in *Lc-tra*.

et al., 2013; Koch *et al*., 2014). The *tra* homologue in honeybees (*Ap. mellifera*, *Apis cerana*, *Apis dorsata*) is called *feminizer* (*fem*) and is duplicated. This paralogue contains a hypervariable region and was identified as the *complementary sex determiner* (*csd*) locus (Hasselmann *et al*., 2008a, 2008b). The hypervariable region is not present in *tra* paralogues of bumblebees and ants, and also does not appear in *Lc-traB*. The similarity of *Lc-tra* and *Lc-traB* sequences between reproductive types points to a duplication event after the divergence from other hymenopterans, but before the split between thelytokous and arrhenotokous lineages of *L. clavipes* (Figs 5, 7).

Identification and sequence variation of Lc-tra2

The nucleotide sequence of the *Lc-tra2* coding DNA sequences (CDS) was predicted from isotig C57958, nucleotide position 134 to 958 (GenBank accession: GAXY02014083) (Peters *et al*., 2017). This translates into a 275-amino-acid sequence containing two RS domains flanking the RBD, consistent with previously identified *tra2* orthologues.

Alternative splicing was detected in *Lc-tra2*, but the alternative splice variants are not sex-specific (Fig. 8). Splicing variation in exon 2 results in a different length of the N-terminal RS domain. The splicing variation at the last three exons results in highly conserved peptides compared to *N. vitripennis* and *Ap. mellifera* (Nissen *et al*., 2012; Geuverink *et al*., 2017). *Lc-tra2*^{A} and *Lc-tra2*^{B} translate at the 3' end of the coding region to a FESRGIG motif, whereas *Lc-tra2*^{C} translates into RY immediately followed by a STOP codon through the inclusion of exon 4. Owing to an A-rich region in the 3' region it was impossible to obtain 3'RACE PCR fragments of *Lc-tra2*. Thus, a splice variant with a poly-A tail at the end of exon 4 could exist, but would not yield a different protein sequence or UTR compared to the variants reported here. *Lc-tra*^{C} is visible in the female samples of Fig. 4,

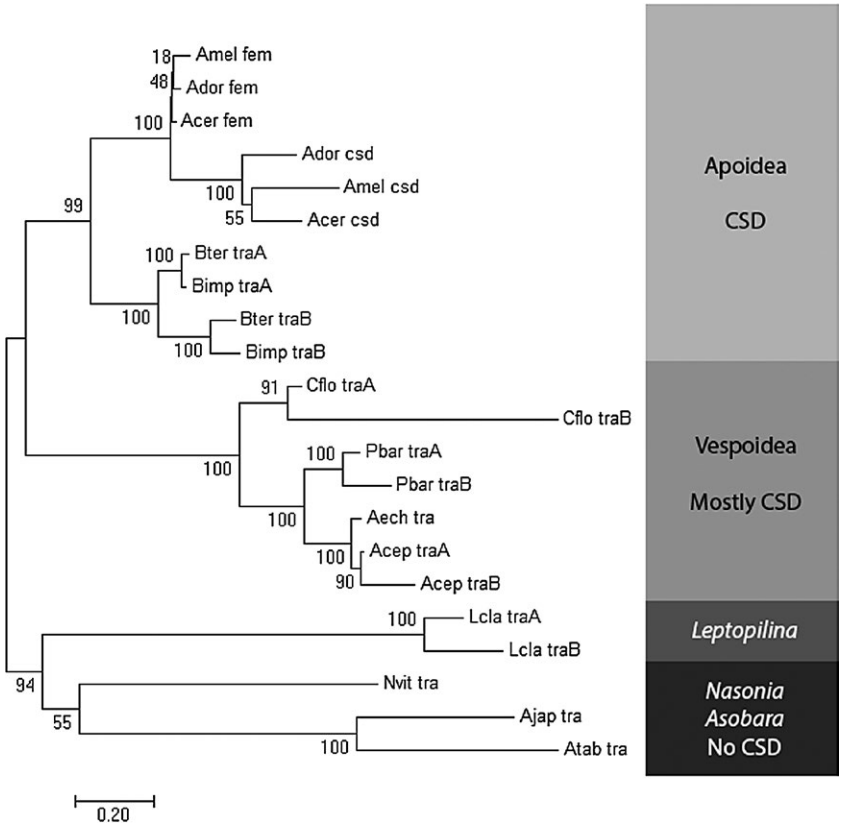

Figure 7. Gene tree of Transformer/Feminizer (TRA/FEM) and the duplications TransformerB/Complementary sex determiner (TRAB/CSD) in Hymenoptera. The analysis is based on predicted protein sequences of the female splice variants (including putative autoregulatory *Ceratitis-Apis-Musca* domain, arginine/serine-rich region and protein-rich region). All 578 positions of the 22 amino acid sequences were used to construct the tree with the maximum likelihood method based on the Jones *et al*. w/freq. model with gamma distribution (Jones et al., 1992). Bootstrap values (1000 replicates) are shown on the branches. The scale bar shows the number of substitutions per site. Amel, *Apis mellifera*; Ador, *Apis dorsata*; Acer, *Apis cerana*; Bter, *Bombus terrestris*; Bimp, *Bombus impatiens*; Cflo, *Camponotus floridanus*; Pbar, *Pogonomyrmex barbatus*; Aech, *Acromyrmex echinatior*; Acep, *Atta cephalotus*; Nvit, *Nasonia vitripennis*; Atab, *Asobara tabida*; Ajap, *Asobara japonica*; Lcla, *Leptopilina clavipes*. CSD absence or presence is noted on the right.

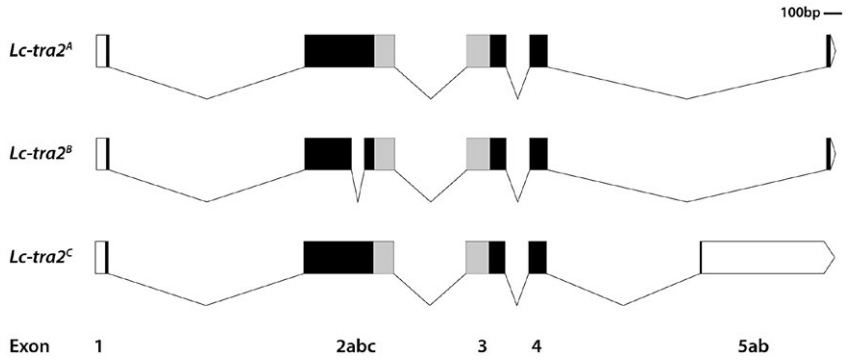

Figure 8. Exon–intron structure of the splice variants of *Leptopilina clavipes transformer 2* (*Lc-tra2*). White boxes represent the 5' and 3' untranslated regions, the black boxes contain the coding sequence. Grey boxes depict the RNA binding domain.

whereas an alternative faint amplicon is visible in the males, but based on various RT-PCRs in this 3' region these transcripts do not seem sex-specific, and are low in abundance. No splicing variation was found between reproductive modes.

Comparison of the genomic region of *Lc-tra2* between the two thelytokous and two arrhenotokous strains yielded a limited number of SNPs (Table 1). None of these SNPs is located on exons in the coding region, confirming the high level of *tra2* conservation between the reproductive types. No sequence variation is present within the reproductive types.

Protein interactions between Lc-tra*,* Lc-traB *and* Lc-tra2

Protein–protein interaction between TRA and TRA2 is required to regulate female-specific splicing of *dsx* pre-mRNA in *Drosophila melanogaster* (Amrein *et al*., 1994), but *in vitro* tests of this molecular interaction have not been performed for the TRA and TRA2 homologues of other insects. We examined the *N. vitripennis* model to assess conservation of this protein interaction in Hymenoptera, using the Yeast 2-Hybrid protein interaction assay (Fields and Song, 1989). Yeast 2-Hybrid vectors containing either a DNA-binding or a transcriptional activation domain were fused to full-length CDS of *N. vitripennis tra* (NV-TRA) and *N. vitripennis tra2* (NV-TRA2). Yeast 2-Hybrid assays demonstrated a weak interaction between NV-TRA and NV-TRA2 (Table 2). NV-TRA, as well as NV-TRA2, interacts with itself, a feature also observed in *Drosophila* (Amrein *et al*., 1994; Table 2).

Subsequently, full-length CDS of *Lc-tra2*^{A}, *Lc-tra*^{F} and *Lc-traB* were cloned in the Yeast 2-hybrid vectors. *Lc-tra2*^{A} (LC-TRA2) was selected based on transcript abundance in both female and male wasps and *Lc-tra*^{F} (LC-TRA) constitutes the only *tra* splice variant that codes for a full-length ORF.

As the results in Table 2 demonstrate, LC-TRA interacts with LC-TRA2 and with TRA2 of *N. vitripennis* (NV-TRA2).

Table 2. Protein–protein interactions in the Yeast 2-Hybrid system. Protein–protein interaction was assessed using the Matchmaker Gal4 Two-Hybrid System 3 as provided by Clontech. Briefly, the first interacting protein is fused to a DNA binding domain for the GAL promoter by cloning in one plasmid, and the second interacting protein is fused to a transcriptional activation domain by cloning in another plasmid. Both plasmids are then expressed together in a yeast strain containing several GAL promoter-driven reporter genes. The DNA binding domain recognizes, and binds, the GAL promoter. If protein interaction between the two cloned proteins occurs, the activation domain is now recruited to the GAL promoter region and induces expression of reporter genes. These reporter genes are typically for auxotrophic markers (in these experiments histidine and adenine), which allows survival of the yeast on 'drop-out' media. Growth is thus a measure of protein–protein interaction of the cloned genes

Insert 1 (binding domain)	Insert 2 (activation domain)	Interaction
Leptopilina TRA	Leptopilina TRA2	++
Leptopilina TRAB	Leptopilina TRA2	++
Leptopilina TRAB	Leptopilina TRA	+
Leptopilina TRA	Leptopilina TRAB	+
Leptopilina TRA	Leptopilina TRA	-
Leptopilina TRAB	Leptopilina TRAB	++
Leptopilina TRA2	Leptopilina TRA2	++
Nasonia TRA	Leptopilina TRA2	++
Leptopilina TRA	Nasonia TRA2	++
Leptopilina TRAB	Nasonia TRA2	++
Nasonia TRA	Leptopilina TRA	-
Nasonia TRA	Nasonia TRA	+
Nasonia TRA	Nasonia TRA2	+
Nasonia TRA2	Nasonia TRA2	++
p53 (negative control)	Leptopilina TRA2	-
Lam (negative control)	Leptopilina TRA2	-
p53 (negative control)	Nasonia TRA2	-
Lam (negative control)	Nasonia TRA2	-

[a]TRA, transformer; TRAB, transformer paralogue.

Conversely NV-TRA interacts with LC-TRA2. This shows the conserved ability of TRA2 to bind diverged TRA homologues. In *D. melanogaster* TRA interacts with itself (see above), but we only observed this for NV-TRA and not for LC-TRA (Table 2). LC-TRAB showed interactions with LC-TRA2 and NV-TRA2, confirming TRA2 binding recognition of TRA-like sequences (Table 2). LC-TRAB also interacts with LC-TRA (Table 2) allowing the possibility of a trimeric protein complex of LC-TRA, LC-TRAB and LC-TRA2. LC-TRAB additionally interacts with itself.

Table 3. Ploidy assessments of arrhenotokous males, arrhenotokous females, thelytokous males and thelytokous females

Strain	Reproductive mode	Sex	Wolbachia infection	Sample size	Ploidy
CA1	Arrhenotoky	Male	No	5	Haploid
CA1	Arrhenotoky	Female	No	4	Diploid
LS1	Thelytoky	Male	No	6	Haploid
LS1	Thelytoky	Female	No	3	Diploid
LS1	Thelytoky	Female	Yes	5	Diploid

As most test combinations yielded an interaction we verified the likelihood of false positive interactions through inclusion of control plasmids that contained nonrelated proteins (murine p53 or human Lamin C). The *L. clavipes* constructs did not interact with these constructs, or with empty constructs (which only express either the binding or the activation domain), indicating that the observed interactions between the sex determination genes are specific (Table 2 and Supporting Information Table S2).

Ploidy of arrhenotokous and thelytokous L. clavipes

Males and females were compared for ploidy between arrhenotokous and thelytokous strains. No differences were detected between sexes of the different reproductive modes. All females were diploid and all males were haploid (Table 3).

Discussion

The transducing level of sex determination is conserved in L. clavipes

Two homologues of *tra* were detected in *L. clavipes* and both displayed strong amino acid sequence conservation compared to hymenopteran TRA orthologues. *Lc-tra* is probably the *tra* orthologue based on the sex-specific splicing of its transcripts, sequence conservation within reproductive type and the high conservation of domains and structure of the LC-TRA protein. It retains all elements required for a conserved sex determination function. Comparison of the arrhenotokous and thelytokous *Lc-tra*^{F} mRNAs and peptides did not reveal much divergence. Genes that have become redundant in thelytokous wasps have been observed to decay (Kraaijeveld *et al*., 2016). In the thelytokous (all-female) lineage this decay was not observed in the genomic region of *Lc-tra* specifying the male-specific exon containing the premature stop codon. Haploid males that are produced by thelytokous females after antibiotic treatment contain the same male-specific splice variants as arrhenotokous haploid males. However, *Lc-tra*^{M} transcripts are not present in thelytokous females infected with *Wolbachia* endosymbionts. This is a notable difference compared to arrhenotokous females, which, beyond their expression of *Lc-tra*^{F}, also express abundant male-specific *Lc-tra* transcripts. Apparently, *Wolbachia* infection prevents the generation of male specific splice forms from *Lc-tra* in females. Yet this entire genomic region of *Lc-tra* is highly conserved between arrhenotokous and thelytokous wasps, and production of male splice forms is possible upon removal of *Wolbachia* endosymbionts. This suggests that the functionality of *tra* is retained in either reproductive mode.

Lc-tra2 contains all conserved regions associated with *tra2*. Its splicing variation at the 3'end corresponds with known variants in *Ap. mellifera*, *N. vitripennis* and *As. tabida* (Nissen *et al*., 2012; Geuverink *et al*., 2017, 2018). We performed protein–protein interaction assays in *N. vitripennis* to test the hypothetical TRA/TRA2 binding complex in Hymenoptera. Both *tra* and *tra2* in *N. vitripennis* are required for the splicing of *tra* pre-mRNA as well as *dsx* pre-mRNA (Verhulst *et al*., 2010a; Geuverink *et al*., 2017). In this study we demonstrated an interaction between TRA and TRA2 in *N. vitripennis*. The conservation of this interaction in *L. clavipes,* combined with the cross interaction of TRA and TRA2 between *L. clavipes* and *N. vitripennis*, enables the possibility that *tra* and *tra2* are also involved in *tra* and *dsx* pre-mRNA splicing in this species.

Involvement of Lc-traB *in sex determination of* L. clavipes

Lc-traB is neither sex-specifically nor alternatively spliced. The absence of the corresponding male-specific exon region of *Lc-tra* results in a default splicing of *Lc-traB* transcripts, similar to the female-specific *tra*. This suggests that *traB* does not need autoregulation of its splicing once switched on. The *Lc-traB* sequence coding for the CAM region associated with autoregulation (Hediger *et al*., 2010) is distinctly different compared to *Lc-tra* (Fig. 3). If *traB* had become obsolete in thelytokous sex determination this would be visible in sequence degeneration. However, this is not observed in any of the four (re)sequenced lineages. *Lc-traB* does group into three haplotypes independent of reproductive mode. This could reflect the evolutionary history of the lineages, rather than a functional implication, as the clustering matches divergence patterns observed with neutral and mitochondrial markers (Kraaijeveld *et al*., 2011). The strong sequence conservation and lack of degeneration in thelytokous lineages suggest a

conserved function. Yet, the lack of male-specific region and default splicing indicate a function different from *Lc-tra*. As LC-TRAB interacts with LC-TRA and LC-TRA2, this function may still be a part of the process of sex determination. Thus, we propose that a combination of *Lc-tra*, *Lc-traB* and *Lc-tra2* may be required for female development in *L. clavipes*.

Implications for the sex determination mechanism of L. clavipes

One of the widespread mechanisms of sex determination in Hymenoptera is CSD. Under CSD female development ensues when one (single-locus CSD) or multiple (multi-locus CSD) loci are heterozygous. The only *csd* locus identified thus far is in *Ap. mellifera* and it is a paralogue of *fem,* the *Ap. mellifera* orthologue of *tra* (Hasselmann *et al*., 2008a). Other paralogues of *tra* have been found in species with CSD, but little is known about their functionality (Schmieder *et al*., 2012; Privman *et al*., 2013). This association potentially reflects a bias in study effort, rather than a true link to CSD. Recently, three homologues of *tra* were reported from the fig wasp *Ceratosolen solmsi*, a species belonging to the Chalcidoidea in which CSD appears absent (van Wilgenburg *et al*., 2006; Heimpel and de Boer, 2008; Jia *et al*., 2016). Transcripts of the two duplicates in *C. solmsi* are only detected in females, but their possible role in sex determination remains unknown. Hence, the presence of a *tra* paralogue is not informative about the presence or absence of CSD.

The lack of evidence for CSD in the *Leptopilina* genus (Biémont and Bouletreau, 1980; Hey and Gargiulo, 1985) requires consideration of the only other reported sex determination mechanism in Hymenoptera: maternal effect genomic imprinting. This has been described for the wasp *N. vitripennis* (Beukeboom and van de Zande, 2010; Verhulst *et al*., 2010a, 2013) and consists of a maternally imprinted (inactivated) sex determination gene [the putative *womanizer* (*wom*) *gene*] that can perform a feminizing function in the zygote. The non-inactivated *wom* of paternal origin in fertilized eggs acts in combination with maternal provisioning of *tra* and *tra2* mRNA to effectuate female development. It is not known if this mechanism, which requires sex determination gene transcripts to be maternally provided to the eggs and involves a paternally provided factor in the fertilized egg, is present in other groups. The presence of two *tra* homologues in *L. clavipes* provides multiple options for maternal effect genes. Additional studies are required to elucidate the thelytokous (uniparental, all-female) mode of sex determination, as under thelytoky a paternally provided element is impossible. How can female development be activated in a zygote containing only maternally provided chromosome sets and gene products? One intriguing possibility is that *Wolbachia* provides this signal. *Wolbachia* may directly interfere with the splicing regulation of *Lc-tra* itself, resulting in the fixed splicing pattern observed in thelytokous adult females. Transcriptomes of early developmental stages need to be procured to identify these signals. This may shed more light on the diversity of sex determination mechanisms in hymenopteran insects and open the possibility of testing endosymbiont interference in insect sex determination.

Experimental procedures

Source material

Stored samples of 12 thelytokous strains (AR1, AR2a, AR3a, Aust, BB1, CDB1a, GBW, KBH, MGS4, STP, WB1a, WB3) and nine arrhenotokous strains (CBY, DC, EJ, EPG, Mol, MS, PdA, PlB, TL), as described in Pannebakker *et al*. (2004b, Kraaijeveld *et al*. (2011) and Table 1, were used to screen divergence of the *tra* genes. The additional SCA strain was collected in Santa Cristina d'Aro (Spain) in October 2015. The wasps were cultured on second-instar *Drosophila phalerata* host larvae at 25 °C under constant light. Individuals from the KBH strain were kindly provided by Todd Schlenke. Females of the KBH strain were cured from their *Wolbachia* infection by feeding honey with 0.5% rifampicin (Schidlo *et al*., 2002); this results in haploid eggs that develop into males (referred to as 'thelytokous males').

Identification of tra *homologues and structure of* tra *in* L. clavipes

Scaffolds containing putative *tra* homologues were identified from the *L. clavipes* genome assembly (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 thelytokous strain GBW were collected from laboratory cultures that were terminated immediately afterwards. RNA extractions were performed with TriZol according to the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). All isolated total RNA was primed with oligo(dT) and random hexamers (in a mixture of 1:6) and reverse transcribed with a RevertAid^{TM} H Minus First Strand cDNA Synthesis Kit (Fermentas, Hanover, MD, USA). Reverse transcription for 3'RACE adapter synthesis was also performed with a RevertAid^{TM} H Minus First Strand cDNA Synthesis Kit (Fermentas) using all isolated total RNA primed with a 3'RACE adapter (5'-GCG AGC ACA GAA TTA ATA CGA CTC ACT ATA GGT 12VN-3'). A 5'RACE adapter containing cDNA was produced according to the manufacturer's instructions (FirstChoice RLM-RACE kit, Ambion, Austin, TX, USA). Sequences of all primers used in this study are shown in Table 4. To assess the *Lc-tra* splice variants present in adult males and females 5'RACE-PCR was performed with outer primer Lcla_tra_5RACE1

Table 4. Overview of primers used in this study

Primer name	Gene/construct	Application	Primer sequence 5'–3'
Lcla_tra_5RACE1	*Lc-tra*	RACE-PCR	ATTGACAAGAGAAGAGAAGC
Lcla_tra_5RACE2	*Lc-tra*	RACE-PCR	CCAGATATGTTTCGGTGAAT
Lcla_tra_3RACE1	*Lc-tra*	RACE-PCR	TGAACCTTTGTTTCGTGGAC
Lcla_tra_3RACE2	*Lc-tra*	RACE-PCR	AACATATCTGGACCCGTCGA
pGEM-T_F	pGEM-T vector	Colony PCR	GTAAAACGACGGCCAGT
pGEM-T_R	pGEM-T vector	Colony PCR	GGAAACAGCTATGACCATG
Lcla_traB_frontF	*Lc-traB*	RT-PCR	GAGACAAGAGAAAAGAAGC
Lcla_traAB_endR	*Lc-traB*	RT-PCR	TGGATTCATGTATCTAGGTGGA
Lcla_tra_spliceA_F	*Lc-tra*	RT-PCR	CAGTCAGAGACAGACGATCC
Lcla_tra_spliceA_R	*Lc-tra*	RT-PCR	TACTTCTCGATGTTCACCTTCC
Lcla_spliceB_F	*Lc-traB*	RT-PCR/PCR	GGACCGAGTACATCATTGAG
Lcla_spliceB_R	*Lc-traB*	RT-PCR/PCR	CGTTGACTTCTCCATTGAATCTG
Lcla_traAB_F	*Lc-tra*/*Lc-traB*	PCR	GTCCATCATTCAGAGACAGAC
Lcla_traA_R	*Lc-tra*	PCR	AGGTCATTATTTATATCGACGG
Lcla_traB_R	*Lc-traB*	PCR	AGGTCATTATTTACAATGATGG
Y2H_Res_Lcla_TraA_F	*Lc-tra*	RT-PCR	CATGGAGGCCGAATTCATGAGACGAAGATCACCTAGTGCA
Y2H_Res_Lcla_TraA_R	*Lc-tra*	RT-PCR	GCAGGTCGACGGATCCCTAAAACTGACGACTGAAACGAGG
Y2H_Res_Lcla_TraB_F	*Lc-traB*	RT-PCR	CATGGAGGCCGAATTCATGAGACGAAGATCTCCTGGTCCA
Y2H_Res_Lcla_TraB_R	*Lc-traB*	RT-PCR	GCAGGTCGACGGATCCCTATTTATGATTAGGAGGAAATCT
Y2H_Lcla_Tra2_NdeI_F	*Lc-tra2*	RT-PCR	AGATTACGCTCATATGGATATGGAGAGGAGTGGAAGTCG
Y2H_Lcla_Tra2_BamHI_R	*Lc-tra2*	RT-PCR	GCAGGTCGACGGATCCTCCAATACCTCTTGATTCAAA
Y2H_Res_Nvit-tra_F	*Nv-tra*	RT-PCR	GTCTTAGAGTAAACATTGGGCTG
Y2H_Res_Nvit-tra_R	*Nv-tra*	RT-PCR	TTTATAGTCCTACGTCGACCT
Y2H_Res_Nvit-tra2_F	*Nv-tra2*	RT-PCR	TCTTTGTTCTACTTTCTCTGTCCC
Y2H_Res_Nvit-tra2_R	*Nv-tra2*	RT-PCR	GAACTTCAACTTTCAGCAACC
Y2H_T7promotor_F	pGBKT7 / pGADT7	Colony PCR	TAATACGACTCACTATAGGGC
Y2H_3'DNA-BD_R	pGBKT7	Colony PCR	TTTTCGTTTTAAAACCTAAGAGTC
Y2H_3'AD_R	pGADT7	Colony PCR	AGATGGTGCACGATGCACAG
Lcla_tra2_5RACE1	*Lc-tra2*	RACE-PCR	GAGAACGACGATAACTTCCTCTG
Lcla_tra2_5RACE2	*Lc-tra2*	RACE-PCR	TGCGACTCCGATGACCACCA
Lcla_tra2_3RACE1	*Lc-tra2*	RACE-PCR	CCAAGACACGAAGATCAAGAGG
Lcla_tra2_3RACE2	*Lc-tra2*	RACE-PCR	AATTGTCGCTATTCTCACCTTTCC
Lcla_tra2_F1	*Lc-tra2*	RT-PCR	GAGTGATATGGAGAGGAGTGGA
Lcla_tra2_R1	*Lc-tra2*	RT-PCR	GGCAACTTTAGCATCTTCAGG
Lcla_tra2_F2	*Lc-tra2*	RT-PCR	TTCACGTTCACGATCAAGGA
Lcla_tra2_R2	*Lc-tra2*	RT-PCR	GTATTCCAACCTTTACATCGTGG
Lcla_tra2_F3	*Lc-tra2*	RT-PCR	CACCTGAAGATGCTAAAGTTGCCA
Lcla_tra2_R3	*Lc-tra2*	RT-PCR	CTTCCCTCTATCATCCAATACCT

and inner primer Lcla_tra_5RACE2 in a reaction at 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 54 °C for 30 s and 72 °C for 60 s, with a final extension of 10 min at 72 °C. Outer primer Lcla_tra_3RACE1 and inner primer Lcla_tra_3RACE2 were used in 3'RACE-PCR in a reaction with DreamTaq (Fermentas). Cycling conditions were 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 59 °C for 30 s and 72 °C for 2 min, with a final extension of 7 min at 72 °C. Resulting PCR fragments were visualized on ethidium bromide-containing 1.5% agarose gel with 1x TAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA).

All RACE-PCR products were ligated into pGEM-T vector (Promega, Madison, WI, USA) after purification using a GeneJET Gel Purification Kit (Fermentas). Ligation products were used to transform competent JM-109 *Escherichia coli* (Promega). Colony PCR was conducted by use of pGEM-T primers at 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 2 min, with a final extension of 7 min at 72 °C.

As the lowered specificity of RACE-PCRs (one gene-specific primer per PCR) rarely permitted the detection of *Lc-traB*, RT-PCRs were used to detect splice variation in this gene. These PCRs were performed with primers Lcla_traB_frontF and Lcla_traAB_endR in a reaction at 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 2 min, with a final extension of 7 min at 72 °C.

PCR-fragments were sequenced on an ABI 3730XL (Applied Biosystems, Foster City, CA, USA) and reads were inspected in Chromas (Technelysium, South Brisbane, Australia) and aligned in MEGA7 (Kumar *et al*., 2016). Exon–intron structure of the genes was constructed by comparing the mRNA sequences to the genomic assembly scaffolds (*Lc-tra*: scf7180005166757, *Lc-traB*: scf7180005164248) and visualized with EXON-INTRON GRAPHIC MAKER (https://wormweb.org/exonintron). Transcript sequences were deposited in GenBank (accession numbers: MG963997–MG964000).

Differential splicing of Lc-tra *and* Lc-traB *in arrhenotokous and thelytokous* L. clavipes

RNA extractions of male and female wasps were performed with TriZol (Invitrogen) according to the manufacturer's protocol. Adult males and females of the

arrhenotokous strain SCA and thelytokous strain KBH were collected from laboratory cultures. RNA extractions were performed as described above.

The presence of sex-specific splice variants of *Lc-tra* in adults was tested with primers Lcla_tra_spliceA_F and Lcla_tra_spliceA_R. Transcripts of *Lc-traB* were detected with primers Lcla_spliceB_F and Lcla_spliceB_R. The cycling-conditions were 94 °C for 3 min, 45 cycles of 94 °C for 30 s, 57 °C (*tra*)/55 °C (*traB*) for 30 s and 72 °C for 2 min, with a final extension of 7 min at 72 °C. The resulting fragments of each category were sequenced to verify their identity as *Lc-tra* male- and female-specific splice variants and *Lc-traB*.

Sequence divergence of Lc-tra *and* Lc-traB *in thelytokous and arrhenotokous individuals*

To assess variation in the *tra* genes between different populations of *L. clavipes* DNA was individually extracted from five females per strain with a standard high salt protocol (Aljanabi and Martinez, 1997). Population variation was tested with the primer Lcla_traAB_F (5'-GTCCATCATTCAGAGACAGAC-3') in combination with Lcla_traA_R (5'-AGGTCATTATTTATATCGACGG-3') and Lcla_traB_R (5'-AGGTCATTATTTACAATGATGG-3'). Reaction conditions were 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 45 s, with a final extension of 7 min at 72 °C. Fragments were sequenced, inspected in Chromas (Technelysium) and aligned in MEGA7 (Kumar *et al*., 2016). A median-joining haplotype network was constructed with PopART (http://popart.otago.ac.nz).

Whole genome sequencing was performed on DNA of three *L. clavipes* lineages (EJ, PdA and MGS4) that had been stored in 96% ethanol at 4 °C. DNA was extracted from groups of 20 females following the animal tissue protocol with spin-columns from a DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA). In short, wasps were air-dried and crushed in 180 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA) in a 1.5-ml microcentrifuge tube using a pestle. After adding 20 µl Proteinase K (600 mAU/µl), the samples were incubated for 90–120 min at 56 °C with frequent vortexing. 5 µl RNAse (100 mg/ml) was added prior to washing the spin-columns according to the standard protocol. The DNA was eluted in 100 µl DNAse-free MilliQ (EMD Millipore, Burlington, Massachusetts, United States) water. Length and integrity of the DNA molecules were checked on a 2100 Bioanalyzer lab-on-chip with a High Sensitivity DNA kit (Agilent, Santa Clara, CA, USA) and purity was assessed using a Nanodrop spectophotometer (Thermo Fisher Scientific Waltham, Massachusetts, United States).

For Illumina library preparation, each DNA sample was fragmented and size-selected to 350 bp using a KAPA HyperPlus Library Preparation Kit (KAPA Biosystems, Boston, MA, USA) according to the supplied protocol (KR1145v215-1). Size range and concentration were assessed using a 2100 Bioanalyzer lab-on-chip with a High Sensitivity DNA kit (Agilent). Library concentration and correct adaptor ligation were assessed using quantitative (q) PCR according to basic protocol 5 documented in Bronner *et al*. (2014). The libraries were stored at –20 °C in 100 µl DNAse-free MilliQ water and sequenced within 4 weeks on an Illumina HiSeq4000 (Illumina, San Diego, California, United States) (150 bp paired-end) at the Leiden Genome Technology Center (Leiden, the Netherlands).

Initial quality checks of the raw reads were conducted using FASTQC (Andrews, 2010). Reads were mapped to the *L. clavipes* reference genome using Bowtie2 (Langmead & Salzberg, 2012). Duplicate reads were removed using PicardTools (https://broadinstitute.github.io/picard/) and indel realignment was conducted using GATK (McKenna *et al*., 2010). Consensus sequences of each strain were constructed by comparing the aligned reads of EJ, PdA and MGS4 to the GBW reference sequence. Full-length amino acid sequences of arrhenotokous and thelytokous LC-TRA and LC-TRAB were aligned with MUSCLE (Edgar, 2004) in Geneious 8 (Biomatters Ltd, Auckland, New Zealand). The intronic regions between *traB* exons 3 and 4 were amplified in DNA samples of the different strains with the same primers as used for RT-PCR: Lcla_spliceB_F

Table 5. Constructs for Yeast 2-Hybrid assay and primers containing restriction sites

Gene	Primer name	Restriction adapters	Annealing temperature (C)
Lc-tra	Y2H_Res_Lcla_TraA_F	*Eco*RI (5')	64
	Y2H_Res_Lcla_TraA_R	*Bam*HI (3')	
Lc-traB	Y2H_Res_Lcla_TraB_F	*Eco*RI (5')	60
	Y2H_Res_Lcla_TraB_R	*Bam*HI (3')	
Lc-tra2	Y2H_Lcla_Tra2_NdeI_F	*Nde*I (5')	62
	Y2H_Lcla_Tra2_BamHI_R	*Bam*HI (3')	
Nv-tra	Y2H_Res_Nvit-tra_F	*Eco*RI (5')	53
	Y2H_Res_Nvit-tra_R	*Bam*HI (3')	
Nv-tra2	Y2H_Res_Nvit-tra2_F	*Eco*RI (5')	56
	Y2H_Res_Nvit-tra2_R	*Bam*HI (3')	

[a]Lc, *Leptopilina clavipes*; Nv, Nasonia vitripennis; tra, transformer; traB, tra paralogue.

and Lcla_spliceB_R. The cycling-conditions were 94 °C for 3 min, 45 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 2 min, with a final extension of 7 min at 72 °C.

Conservation of Lc-tra *and* Lc-traB

The following sequences were used in alignments and the gene tree: *Ap. mellifera fem* (AAS86667) and *csd* (AAS86653), *Ap. dorsata fem* (ABV56232) and *csd* (ABW36165), *Ap. cerana fem* (ABV56230) and *csd* (ABV58877), *Bombus terrestris traA* (NP_001267853) and *traB* (XP_003394693), *Bombus impatiens traA* (XP_003493796) and *traB* (XP_003491525), *N. vitripennis tra* (NP_001128299). Ant protein sequences were obtained from Privman *et al*.'s (2013) supplementary materials. Alignments were produced in CLC WORKBENCH (CLCbio, Aarhus, Denmark).

Identification and sequence variation of Lc-tra2

5'RACE-PCR was performed with outer primer Lcla_tra2_5RACE1 and inner primer Lcla_tra2_5RACE2. The PCR cycles were as follows: 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min, with a final extension of 10 min at 72 °C. Outer primer Lcla_tra2_3RACE1 and inner primer Lcla_tra2_3RACE2 were used in 3'RACE-PCR with the same cycling conditions as the 5'RACE-PCR. RT-PCRs to verify and detect further splice variation were performed with primers Lcla_tra2_F1/Lcla_tra2_R1, Lcla_tra2_F2/Lcla_tra2_R2 and Lcla_tra2_F3/Lcla_tra2_R3 under the following conditions: 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 57 °C(primer set1) /53 °C (primer set2)/52 °C (primer set3) for 30 s and 72 °C for 2 min, with a final extension of 10 min at 72 °C. Purification of PCR products, ligation, transformation, colony PCR and sequencing were performed according to the procedures described for *Lc-tra* identification. Transcript sequences were deposited in GenBank (accession numbers: MG963994–MG963996). *Lc-tra2* consensus sequences of each resequenced strain were constructed by comparing the aligned reads of EJ, PdA and MGS4 to the GBW reference sequence.

Protein interactions between Lc-tra*,* Lc-tra*B and* Lc-tra*2*

To test protein–protein interactions of the sex determination genes, cDNA of *L. clavipes* and *N. vitripennis* was obtained through the methodology described above. The selected transcripts were the female-specific variant of *Lc-tra* and *Nv-tra* (Werren *et al*., 2010) and the single splice variant of *Lc-traB*. The chosen splice variants of *tra2* (*Lc-tra2*^{A} and *Nv-tra2*^{A}) were most abundant in *N. vitripennis* and conserved in other Hymenoptera (translated into FESRGIG motif) (Geuverink *et al*., 2017, 2018). Primers containing adapters that add restriction sites for *Eco*RI/*Nde*I (5'end) and *Bam*HI (3'end) were used to amplify full-length transcripts in RT-PCRs. These primers and corresponding restriction sites are displayed in Table 5 and the primer sequences are given in Table 4. The cycling conditions were 94 °C for 3 min, 45 cycles of 94 °C for 30 s, 53–64 °C for 30 s and 72 °C for 90 s, with a final extension of 7 min at 72 °C. Annealing temperatures per primer pair are shown in Table 4. PCR products were visualized on ethidium-bromide-containing 1.5% agarose gel with 1x TAE buffer and were purified from gel using a GeneJET Gel Extraction Kit (Thermo Fisher Scientific) according to the manufacturer's protocol.

All PCR products were digested with restriction enzymes (Table 5) using a double digestion. Plasmids pGBKT7 and pGADT7 (Clontech, Mountain View, CA, USA) were also double digested with both *Eco*RI/*Bam*HI and *Nde*I/*Bam*HI. Digestion reactions for transcripts Lc-tra, Lc-traB, Nv-tra and Nv-tra2 consisted of: 1 µg cleaned-up PCR product and 1 µg of both plasmids pGBKT7 and pGADT7, 1 µl *Eco*RI, 0.5 µl *Bam*HI and 2 µl *Bam*HI buffer; the volume was increased to 20 µl with MilliQ. These reactions were incubated at 37 °C for 16 h, followed by a 20-min incubation at 80 °C. Digestion reactions for transcript Lc-tra^{2} consisted of: 1 µg cleaned-up PCR product and 1 µg of both plasmids pGBKT7 and pGADT7, 4 µl *Nde*I, 1 µl *Bam*HI and 2 µl *Bam*HI buffer; the volume was made up to 23.5 µl with MilliQ. These reactions were incubated at 37 °C for 16 h, followed by a 20-min incubation at 80 °C. Digested PCR products were ligated into pGBKT7 and pGADT7 using T4 DNA Ligase (New England Biolabs, Beverly, MA, USA) to yield both bait and prey vectors containing the genes listed in Table 5. Control plasmids pGBKT7-53 (murine) and pGBKT7-lam (human) (Clontech) were included to account for the possibility of false positive detections.

Plasmids containing bait and prey constructs were transformed into competent JM-109 *Escherichia coli* (Promega). Colony PCR was performed using primers: Y2H_T7promotor_F and Y2H_3'DNA-BD_R for pGBKT7 constructs and Y2H_T7promotor_F and Y2H_3'AD_R for pGADT7 constructs. The cycling-conditions were 94 °C for 3 min, 45 cycles of 94 °C for 30 s, 50 °C (AD)/55 °C (BD) for 30 s and 72 °C for 90 s, with a final extension of 7 min at 72 °C. PCR products were sequenced to confirm that no PCR errors were generated and that the protein is fused in-frame with the vector promotor. Plasmids containing the correct genes were isolated from the colonies using a GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific).

pGBKT7 vectors containing binding domains were introduced into yeast strain AH109, and pGADT7 vectors containing activation domains were introduced into yeast strain Y187 to test for protein interactions.

All experimental procedures were conducted according to the Matchmaker GAL4 Two-Hybrid System 3 manual (Clontech). Protein interactions were identified by observing the growth of transformants on SD-Ade/-His/-Leu/-Trp plates as a result of the transcription of reporter genes (Fig. S2).

Ploidy of arrhenotokous and thelytokous L. clavipes

Ploidy of arrhenotokous males, arrhenotokous females, thelytokous males and thelytokous females was confirmed by flow cytometry analysis. Newly collected arrhenotokous strain CA1 and thelytokous strain LS1 were used in this assay (Table S1). A thin layer of yeast mixture containing 2.5 mg tetracycline per gram of dry yeast was added to agar bottles. Second-instar *D. phalerata* larvae were added to the bottle and parasitized by thelytokous LS1 females. All emerging F1 offspring were still female, but cured of their *Wolbachia* infection. They were hosted on regular bottles containing second-instar *D. phalerata* host larvae for parasitization. The emerging F2 offspring solely consisted of males. These thelytokous males, their cured thelytokous mothers and nontreated thelytokous females were used to assess ploidy. Adult wasp heads were homogenized in Galbraith buffer (21 mM $MgCl_2$, 30 mM tri-sodium citrate hydrate, 20 mM MOPS, 0,1% Triton X-100, 1 mg/l RNAse A) using a motorized pestle, filtered by 35-µm cell strainer caps (BD Falcon Cell strainer #352235, BD Biosciences, San Jose, CA, USA) and stained with propidium iodide (Sigma, St Louis, MO, USA). Samples were loaded on a MACSQuant Analyzer (Miltenyi Biotec, Bergisch Gladbach, Germany) and analysed with FlowLogic software (Miltenyi Biotec).

Acknowledgements

We thank Rogier Houwerzijl and Peter Hes for assistance with wasp and fly culturing, Todd Schlenke for supplying the KBH strain, students from the minor Genomics at the University of Applied Sciences Leiden for preparing and sequencing the HiSeq libraries, Mieke Geuverink-Bloemen and Paul Geuverink for collecting the LS1 strain and Anna Rensink for collecting the CA1 strain. This research was funded by Netherlands Organisation for Scientific Research grants no. 854.10.001 and no. 824.15.015 to L.W.B. The authors declare no conflict of interest.

References

Aljanabi, S.M. and Martinez I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, **25**(22), 4692–4693.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller W., *et al*. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**(17), 3389–3402.

Amrein, H., Hedley, M.L. and Maniatis, T. (1994) The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by Transformer 2. *Cell*, **76**, 735–746.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Beukeboom, L.W. and Perrin, N. (2014) *The Evolution of Sex Determination*. Oxford: Oxford University Press.

Beukeboom, L.W. and van de Zande, L. (2010) Genetics of sex determination in the haplodiploid wasp *Nasonia vitripennis* (Hymenoptera: Chalcidoidea). *Journal of Genetics*, **89**(3), 333–339.

Beye, M., Hasselmann, M., Fondrk, M.K., Page, R.E. and Omholt, S.W. (2003) The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell*, **114**, 419–429.

Biémont, C. and Bouletreau, M. (1980) Hybridization and inbreeding effects on genome coadaptation in a haplo-diploid hymenoptera: *Cothonaspis boulardi* (Eucoilidae). *Experientia*, **36**, 45–47.

Bopp, D., Saccone, G. and Beye, M. (2014) Sex determination in insects: variations on a common theme. *Sexual Development*, **8**(1-3), 20–28.

Bronner, I.F., Quail, M.A., Turner, D.J., Swerdlow, H., Bronner, I.F., Quail, M.A. *et al*. (2014) Improved protocols for illumina sequencing. *Current Protocols in Human Genetics*. 80, 18.2.1–18.2.42.

Burghardt, G., Hediger, M., Siegenthaler, C., Moser, M., Dübendorfer, A. and Bopp, D. (2005) The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. *Development Genes and Evolution*, **215**(4), 165–176.

Concha, C. and Scott, M.J. (2009) Sexual development in *Lucilia cuprina* (Diptera, Calliphoridae) is controlled by the *transformer* gene. *Genetics*, **182**, 785–798.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Fields, S. and Song, O. (1989) A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246.

Geuverink, E. and Beukeboom, L.W. (2014) Phylogenetic distribution and evolutionary dynamics of the sex determination genes *doublesex* and *transformer* in insects. *Sexual Development*, **8**, 38–49.

Geuverink, E., Rensink, A.H., Rondeel, I., Beukeboom, L.W., van de Zande, L. and Verhulst, E.C. (2017) Maternal provision of *transformer-2* is required for female development and embryo viability in the wasp *Nasonia vitripennis*. *Insect Biochemistry and Molecular Biology*, **90**, 23–33.

Geuverink, E., Verhulst, E.C., van Leussen, M., van de Zande, L. and Beukeboom, L.W. (2018) Maternal provision of non-sex-specific *transformer* messenger RNA in sex determination of the wasp *Asobara tabida*. *Insect Molecular Biology*, **27**, 99–109.

Hasselmann, M., Gempe, T., Schiøtt, M., Nunes-Silva, C.G., Otte, M. and Beye, M. (2008a) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature*, **454**, 519–522.

Hasselmann, M., Vekemans, X., Pflugfelder, J., Koeniger, N., Koeniger, G., Tingek, S. *et al*. (2008b) Evidence for convergent nucleotide evolution and high allelic turnover rates at the complementary sex determiner gene of Western and Asian honeybees. *Molecular Biology and Evolution*, **25**, 696–708.

Hediger, M., Henggeler, C., Meier, N., Perez, R., Saccone, G. and Bopp, D. (2010) Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. *Genetics*, **184**, 155–170.

Hedley, M.L. and Maniatis, T. (1991) Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. *Cell*, **65**, 579–586.

Heimpel, G.E. and de Boer, J.G. (2008) Sex determination in the Hymenoptera. *Annual Review of Entomology*, **53**, 209–230.

Herpin, A. and Schartl, M. (2015) Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. *EMBO reports*, **16**, 1260–1274.

Hey, J. and Gargiulo, M.K. (1985) Sex-ratio changes in *Leptopilina heterotoma* in response to inbreeding. *Journal of Heredity*, **76**, 209–211.

Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H. and Shimura, Y. (1991) Control of *doublesex* alternative splicing by transformer and transformer-2 in *Drosophila*. *Science*, **252**, 833–836.

Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H. and Shimura, Y. (1992) Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. *Proceedings of the National Academy of Sciences*, **89**, 8092–8096.

Jia, L.-Y., Xiao, J.-H., Xiong, T.-L., Niu, L.-M. and Huang, D.-W. (2016) The *transformer* genes in the fig wasp *Ceratosolen solmsi* provide new evidence for duplications independent of complementary sex determination. *Insect Molecular Biology*, **25**, 191–201.

Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*, **8**, 275–282.

Koch, V., Nissen, I., Schmitt, B. and Beye, M. (2014) Independent evolutionary origin of *fem* paralogous genes and complementary sex determination in hymenopteran insects. *PLoS One*, **9**, e91883.

Kraaijeveld, K., Franco, P., de Knijff, P., Stouthamer, R. and van Alphen, J.J.M. (2011) Clonal genetic variation in a *Wolbachia*-infected asexual wasp: horizontal transmission or historical sex? *Molecular Ecology*, **20**, 3644–3652.

Kraaijeveld, K., Anvar, Y., Frank, J., Schmitz, A., Bast, J., Wilbrandt, J. *et al*. (2016) Decay of sexual trait genes in an asexual parasitoid wasp. *Genome Biology and Evolution*, evw273.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870–1874.

Lagos, D., Koukidou, M., Savakis, C. and Komitopoulou, K. (2007) The *transformer* gene in *Bactrocera oleae*: the genetic switch that determines its sex fate. *Insect Molecular Biology*, **16**, 221–230.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359.

Lattorff, H. M.G., Moritz, R.F.A. and Fuchs, S. (2005) A single locus determines thelytokous parthenogenesis of laying honeybee workers (*Apis mellifera capensis*). *Heredity*, **94**, 533–537.

Liu, G., Wu, Q., Li, J., Zhang, G. and Wan, F. (2015) RNAi-mediated knock-down of *transformer* and *transformer 2* to generate male-only progeny in the oriental fruit fly, *Bactrocera dorsalis* (Hendel). *PLoS One*, **10**, e0128892.

Martín, I., Ruiz, M.F. and Sánchez, L. (2011) The gene *transformer-2* of *Sciara* (Diptera, Nematocera) and its effect on *Drosophila* sexual development. *BMC Developmental Biology*, **11**, 19.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A. *et al*. (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297–1303.

Mine, S., Sumitani, M., Aoki, F., Hatakeyama, M. and Suzuki, M.G. (2017) Identification and functional characterization of the sex-determining gene *doublesex* in the sawfly, *Athalia rosae* (Hymenoptera: Tenthredinidae). *Applied Entomology and Zoology*, **52**, 497–509.

Nagoshi, R.N. and Baker, B.S. (1990) Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: cis-acting mutations in exon sequences alter sex-specific RNA splicing patterns. *Genes & Development*, **4**, 89–97.

Nagoshi, R.N., McKeown, M., Burtis, K.C., Belote, J.M. and Baker, B.S. (1988) The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*. *Cell*, **53**, 229–236.

Nissen, I., Müller, M. and Beye, M. (2012) The *Am-tra2* gene is an essential regulator of female splice regulation at two levels of the sex determination hierarchy of the honeybee. *Genetics*, **192**, 1015–1026.

Niu, B.-L., Meng, Z.-Q., Tao, Y.-Z., Lu, S.-L., Weng, H.-B., He, L.-H. *et al*. (2005) Cloning and alternative splicing analysis of *Bombyx mori transformer-2* Gene using Silkworm EST Database. *Acta Biochimica et Biophysica Sinica*, **37**, 728–736.

O'Neill, S.L., Hoffmann, A.A. and Werren, J.H. (1997) *Influential Passengers, Inherited Microorganisms and Arthropod Reproduction*. Oxford: Oxford University Press.

Pane, A., Salvemini, M., Delli Bovi, P., Polito, C. and Saccone, G. (2002) The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development*, **129**, 3715–3725.

Pane, A., de Simone, A., Saccone, G. and Polito, C. (2005) Evolutionary conservation of *Ceratitis capitata transformer* gene function. *Genetics*, **171**, 615–624.

Pannebakker, B.A., Pijnacker, L.P., Zwaan, B.J. and Beukeboom, L.W. (2004a) Cytology of *Wolbachia*-induced parthenogenesis in *Leptopilina clavipes* (Hymenoptera: Figitidae). *Genome*, **47**, 299–303.

Pannebakker, B.A., Zwaan, B.J., Beukeboom, L.W. and van Alphen, J.J.M. (2004b) Genetic diversity and *Wolbachia* infection of the *Drosophila* parasitoid *Leptopilina clavipes* in western Europe. *Molecular Ecology*, **13**, 1119–1128.

Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K. *et al*. (2017) Evolutionary history of the Hymenoptera. *Current Biology*, **24**, 1586–1591.

Privman, E., Wurm, Y. and Keller, L. (2013) Duplication and concerted evolution in a master sex determiner under balancing selection. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 2012–2968.

Ruiz, M.F., Milano, A., Salvemini, M., Eirín-López, J.M., Perondini, A.L.P., Selivon, D. *et al*. (2007) The gene *transformer* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. *PloS One*, **2**, e1239.

Saccone, G., Salvemini, M. and Polito, L.C. (2011) The *transformer* gene of *Ceratitis capitata*: a paradigm for a conserved epigenetic master regulator of sex determination in insects. *Genetica*, **139**, 99–111.

Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L.C. and Saccone, G. (2009) *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. *The International Journal of Developmental Biology*, **53**, 109–120.

Salvemini, M., D'Amato, R., Petrella, V., Aceto, S., Nimmo, D., Neira, M. *et al*. (2013) The orthologue of the fruitfly sex behaviour gene *fruitless* in the mosquito *Aedes aegypti*: evolution of genomic organisation and alternative splicing. *PLoS One*, **8**, e48554.

Sandrock, C. and Vorburger, C. (2011) Single-locus recessive inheritance of asexual reproduction in a parasitoid wasp. *Current Biology*, **21**, 433–437.

Sarno, F., Ruiz, M.F., Eirín-López, J.M., Perondini, A.L.P., Selivon, D. and Sánchez, L. (2010) The gene *transformer-2* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. *BMC Evolutionary Biology*, **10**, 140.

Schetelig, M.F., Milano, A., Saccone, G. and Handler, A.M. (2012) Male only progeny in *Anastrepha suspensa* by RNAi-induced sex reversion of chromosomal females. *Insect Biochemistry and Molecular Biology*, **42**, 51–57.

Schidlo, N.S., Pannebakker, B.A., Zwaan, B.J., Beukeboom, L.W. and van Alphen, J.J.M. (2002) Curing thelytoky in the *Drosophila* parasitoid *Leptopilina clavipes* (Hymenoptera: Figitidae). *Proceedings of the Section Experimental and Applied Entomology*, **13**, 93–96.

Schmieder, S., Colinet, D. and Poirié, M. (2012) Tracing back the nascence of a new sex-determination pathway to the ancestor of bees and ants. *Nature Communications*, **3**, 895.

Shukla, J.N. and Palli, S.R. (2012) Sex determination in beetles: production of all male progeny by parental RNAi knockdown of *transformer*. *Scientific Reports*, **2**, 602.

Shukla, J.N. and Palli, S.R. (2013) *Tribolium castaneum* Transformer-2 regulates sex determination and development in both males and females. *Insect Biochemistry and Molecular Biology*, **43**, 1125–1132.

Stouthamer, R. (1997) *Wolbachia*-induced parthenogenesis. In:O'Neill, S.L., Hoffmann, A.A. and Werren, J.H. (Eds.) *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*. Oxford: Oxford University Press, pp. 102–124.

Tian, M. and Maniatis, T. (1992) Positive control of pre-mRNA splicing in vitro. *Science*, **256**, 237–240.

Tian, M. and Maniatis, T. (1993) A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. *Cell*, **74**, 105–114.

Van Wilgenburg, E., Driessen, G. and Beukeboom, L.W. (2006) Single locus complementary sex determination in Hymenoptera: an "unintelligent" design? *Frontiers in Zoology*, **3**, 1.

Verhulst, E.C., Beukeboom, L.W. and van de Zande, L. (2010a) Maternal control of haplodiploid sex determination in the wasp *Nasonia*. *Science*, **328**, 620–623.

Verhulst, E.C., van de Zande, L. and Beukeboom, L.W. (2010b) Insect sex determination: it all evolves around *transformer*. *Current Opinion in Genetics and Development*, **20**, 376–383.

Verhulst, E.C., Lynch, J.A., Bopp, D., Beukeboom, L.W. and van de Zande, L. (2013) A new component of the *Nasonia* sex determining cascade is maternally silenced and regulates *transformer* expression. *PloS One*, **8**, e63618.

Werren, J.H., Richards, S., Desjardins, C.A., Niehuis, O., Gadau, J., Colbourne, J.K. *et al*. (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science*, **327**, 343–348.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Alignment of female-specific TRA and non-specific TRAB amino acid sequences of strains EJ, PdA, GBW and MGS4 (populations previously described in Kraaijeveld *et al*., 2011). The HYM and CAM domain are depicted on top of the sequences.

Figure S2. Displayed on the left are pictures of Yeast-2 hybrid matings showing growth on SD-Ade/-His/-Leu/-Trp plates (QDO). Tables summarizing each set of matings are displayed on the right.

Table S1. Strains of *L. clavipes* used in this study.

Table S2. Protein-protein interactions in the Yeast 2-Hybrid system. Protein-protein interaction was assessed using the Matchmaker Gal4 Two-Hybrid System 3 as provided by Clontech. Sex determination genes constructed in combination with either the binding domain or the activation domain were tested against empty constructs.