Quantitative Trait Loci involved in the reproductive success of a parasitoid wasp

Romain Benoist, Claire Capdevielle-Dulac, Célina Chantre, Rémi Jeannette, Paul-André Calatayud, Jean-Michel Drezen, Stéphane Dupas, Arnaud Le Rouzic, Bruno Le Ru, Laurence Moreau, et al.

Quantitative Trait Loci involved in the reproductive success of a parasitoid wasp

Romain Benoist[1], Claire Capdevielle-Dulac[1], Célina Chantre[1], Rémi Jeannette[1], Paul-André Calatayud[1,2], Jean-Michel Drezen[3], Stéphane Dupas[1], Arnaud Le Rouzic[1], Bruno Le Ru[1], Laurence Moreau[4], Erwin Van Dijk[5], Laure Kaiser[1], Florence Mougel[1,*]

[1] *Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, 91198, Gif-sur-Yvette, France.*
[2] *icipe, International Center of Insect Physiology and Ecology, P.O. Box 30772-00100, Nairobi, Kenya*
[3] *Institut de Recherche sur la Biologie de l'Insecte (UMR 7261, CNRS, Université Tours), 37200 Tours, France*
[4] *Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, 91190, Gif-sur-Yvette, France.*
[5] *Institut de Biologie Intégrative de la Cellule (UMR 9198, CNRS, CEA, Université Paris-Sud, Université Paris-Saclay), Avenue de la Terrasse, 9190 Gif-sur-Yvette, France*

** Corresponding author: Florence.mougel-imbert@egce.cnrs-gif.fr*

Running Title: Reproductive success QTL of a parasitoid wasp

Abstract

Dissecting the genetic basis of intraspecific variations in life history traits is essential to understand their evolution, notably for potential biocontrol agents. Such variations are observed in the endoparasitoid *Cotesia typhae* (Hymenoptera: Braconidae), specialized on the pest *Sesamia nonagrioides* (Lepidoptera: Noctuidae). Previously, we identified two strains of *C. typhae* that differed significantly for life history traits on an allopatric host population. To investigate the genetic basis underlying these phenotypic differences, we used a Quantitative Trait Locus (QTL) approach based on Restriction Site Associated DNA markers. The characteristic of *C. typhae* reproduction allowed us generating sisters sharing almost the same genetic content, named clonal sibship. Crosses between individuals from the two strains were performed to generate F2 and F8 recombinant CSS. The genotypes of 181 clonal sibships were determined as well as the phenotypes of the corresponding 4000 females. Informative markers were then used to build a high-quality genetic map. These 465 markers spanned a total length of 1300 cM and were organized in 10 linkage groups which corresponded to the number of *C. typhae* chromosomes. Three QTLs were detected for parasitism

success and two for offspring number, while none were identified for sex ratio. The QTLs explained respectively 27.7% and 24.5% of the phenotypic variation observed. The gene content of the genomic intervals was investigated based on the genome of *C. congregata* and revealed 67 interesting candidates, as potentially involved in the studied traits, including components of the venom and of the symbiotic virus (bracovirus) shown to be necessary for parasitism success in related wasps.

Keywords: Quantitative Trait Loci, parasitism success, offspring number, polydnavirus, venom, parasitoid, linkage map

1 Introduction

Biotic interactions exert a strong selection pressure on living organisms and constitute a major evolutionary strength. Among them, intimate relationships such as host-parasite associations may lead to specific evolutionary patterns, namely co-evolution (Thompson, 2009; Woolhouse, Webster, Domingo, Charlesworth, & Levin, 2002). The process of co-evolution corresponds to adaptive changes in the two partners under reciprocal influence. It requires genetic variation within the parasite and host populations influencing the outcome of the interaction. Deciphering the components of this genetic variation, *i.e.* the number of genes involved, the interactions between them, their potential pleiotropic effects and their impact on the organism's fitness, is of prime importance to elucidate the genomic bases of co-adaptation.

Among parasites, insect parasitoids represent an interesting case study of environmental adaptation both to improve our knowledge of biotic interactions and in the applied perspective of biocontrol (Wang, Liu, Shi, Huang, & Chen, 2019). These species spend their larval stage as parasites and live freely at the adult stage. They behave in an intermediate way between parasites and predators because they need to kill their host to develop. The larval environment is thus mainly biotic and parasitoid effective reproduction relies on fine adaptation to their host (Godfray, 1994).

The host adaptation encompasses the ability to detect and parasitize efficiently the host. Host detection is based on localizing the host as well as recognizing its suitability. Once a suitable host is found, parasitoids have to optimize their oviposition behavior to maximize their fitness. This is illustrated by the occurrence of different oviposition strategies (Godfray, 1994; Waage & Greathead, 1986). Usually, parasitoids are classified as solitary or gregarious depending upon the offspring number produced per host, with one for solitary (not necessarily implying the injection of only one egg) and several for gregarious. In haplodiploid Hymenoptera, oviposition strategy also includes sex ratio of the

progeny, which corresponds to the proportion of fertilized eggs that a female lays (Heimpel & de Boer, 2008). Both the number of eggs injected (and therefore the offspring number) and the sex ratio are under adaptive constraints. For example, the hosts' carrying capacity limits the number of parasitoid larvae that could develop and thus impacts the behavior a species will adopt in the distribution of eggs among hosts encountered by a female (Godfray, 1994; Le Masurier, 1987, 1991). Environmental factors are also perceived and taken into account by parasitoid females to adjust their clutch size and in some cases their sex ratio (Charnov & Skinner, 1985; Waage & Ming, 1984). Those two traits (offspring number and sex ratio) are therefore genetically and environmentally determined.

Following oviposition, parasitoid larvae development depends on the parasitoid ability to inhibit host immune defenses and alter its growth to the parasitoid benefit. This is especially important for endoparasitoid species which develop in host haemocoele, where host haemocytes are recruited to encapsulate parasitoid eggs or larvae (Quicke, 2014). The virulence arsenal developed by parasitoids to ensure the development of their progeny, hereafter defined as parasitism success, is quite phenomenal and combines factors derived from maternal and embryonic origins. Depending on species, the maternal factors encompass venom, ovarian proteins, polydnavirus and virus-like particles and are injected along with eggs in the host (Asgari & Rivers, 2011; Herniou et al., 2013; Pennacchio & Strand, 2006; Pichon et al., 2015; Strand & Burke, 2015). The embryonic factors encompass teratocytes, which are cells derived from the membrane surrounding the parasitoid embryo and parasitoid larvae itself (Strand, 2014). The most studied factors are probably venom and polydnavirus. Venom is mainly devoted to host metabolism manipulation, ensuring parasitoid development (Mrinalini & Werren, 2016). Within hymenoptera the venom gland is a conserved organ but venom composition seems to be highly variable depending on species (Moreau & Asgari, 2015). Polydnaviruses are double-stranded DNA viruses associated with few groups of hymenopteran parasitoids (Webb, 1998). The polydnavirus associated with Braconidae are called Bracoviruses,

93 whereas those associated with Ichneumonidae are called Ichnoviruses. They derive from the
94 integration viruses into the genome of ancestral parasitoids and are now vertically transmitted (Bézier,
95 Annaheim, et al., 2009; Bézier, Herbinière, Lanzrein, & Drezen, 2009; Drezen et al., 2017; Volkoff et al.,
96 2010). In species harboring polydnaviruses, viral particles contain viral DNA circles bearing virulence
97 genes which are expressed in host tissues and involved in the inactivation of the host immune system
98 and alteration of host growth (Beckage & Drezen, 2012; Beckage & Gelman, 2004; Edson, Vinson,
99 Stoltz, & Summers, 1981; Marti, 2003; Wyler & Lanzrein, 2003).

100 Besides the interspecific variation in virulence arsenal and more generally in reproductive success
101 traits, intraspecific divergence in life history traits is also observed in parasitoid species (Chassain &
102 Bouletreau, 1987; Dubuffet et al., 2009; Henter, 1995; Kaiser, Couty, & Perez-Maluf, 2009; Legner,
103 1987; Orzack & Gladstone, 1994). This polymorphism underlies the ability of parasitoids to evolve in
104 response to host selection. Identifying the polymorphic genes involved in parasitoid success is thus of
105 prime importance to identify key components and understand the physiological and behavioral basis
106 of host adaptation.

107 The identification of genes underlying such complex traits is hampered by their multifactorial
108 determinism and high plasticity. Quantitative trait loci (QTL) approaches are optimal to overcome such
109 complex situations (Broman & Sen, 2009; Lander & Botstein, 1989; Mackay, Stone, & Ayroles, 2009).
110 These approaches are also particularly relevant in Hymenopteran species because of their haplodiploid
111 sex determination system. Indeed, in these species males are haploid and produce spermatozoids all
112 bearing identical genetic content. So, when a male is crossed with a completely homozygous female,
113 all daughters produced are genetically identical. This allows us to repeat phenotypic measures on
114 females with the same genotype, reducing the impact of environmental variation (Dupas, Frey, &
115 Carton, 1998; Pannebakker, Watt, Knott, West, & Shuker, 2011; Velthuis, Yang, Van Opijnen, & Werren,
116 2005). To our knowledge, in parasitoids, the QTL approach has been performed only once on sex ratio

and offspring number in *Nasonia vitripennis* (Pannebakker et al., 2011) and never performed to study parasitism success. In this paper we developed a QTL approach to determine the genetic basis of these traits in a parasitoid species: *Cotesia typhae* (Fernández-Triana) (Hymenoptera, Braconidae).

Cotesia typhae is an African gregarious endoparasitoid, parasitizing exclusively larvae of the crop pest *Sesamia nonagrioides* (Lefebvre) (Lepidoptera, Noctuidae). It was formerly undistinguished from the generalist species *Cotesia sesamiae* composed of several populations with different host ranges, but was recently recognized as a distinct species (Kaiser, Dupas, et al., 2017; Kaiser, Fernandez-Triana, et al., 2017; Kaiser et al., 2015). The ability of *C. typhae* to parasitize a population of *S. nonagrioides,* an invasive species in France, was previously studied in a biological control perspective (Benoist et al., 2017). This study highlighted differences between two strains of *C. typhae* named Kobodo and Makindu specifically in their reproductive success on this allopatric host population. Indeed, Kobodo females had a higher rate of parasitism success and produced more offsprings than Makindu females under laboratory conditions. The marked difference in the reproductive success of the two strains sets the stage for studying the genetic basis of these traits. Furthermore, an annotated genome of the close relative *C. congregata* recently became available, which can be used as a reference for the identification of genes within QTL (Gauthier et al., 2020).

The aim of this work was to decipher the genetic architecture involved in the variation of the reproductive success of *C. typhae*. To do this, we developed a QTL approach based on Restriction site Associated DNA markers (RAD-tags) to detect genomic regions associated with parasitism success, offspring number and sex ratio traits. The gene content in the identified QTL region was investigated based on the annotated genome of *Cotesia congregata*.

2 Materials and Methods

2.1 Biological Material

The Kobodo and Makindu *C. typhae* parasitoid strains were obtained from adults emerged from naturally parasitized *S. nonagrioides* caterpillars collected in the field at two localities in Kenya: Kobodo (0.679S, 34.412E; West Kenya; 3 caterpillars collected in 2013) and Makindu (2.278S, 37.825E; South-East Kenya; 10 caterpillars collected in 2010-2011). These strains were reared separately at the International Centre of Insect Physiology and Ecology (ICIPE, Nairobi, Kenya). Isofemale lines were produced and maintained from these rearings from 2015 at the Evolution, Génome, Comportement et Ecologie laboratory (EGCE, Gif-sur-Yvette, France), where cross experiments and phenotyping were performed.

Two host strains of *S. nonagrioides* were used: a Kenyan strain initiated from caterpillars collected at Makindu and a French strain initiated from individuals collected in maize fields (Longage-Berat area in Haute-Garonne district, 43.368N, 1.192E and within a 10 km distance) which was renewed yearly. The French strain was used for phenotyping, whereas the Kenyan strain was used for *C. typhae* rearing to prevent any adaptation of the parasitoid to the French host.

The rearing protocol of *C. typhae* and *S. nonagrioides* is detailed in Benoist et al. (2020).

2.2 Genetic cross-design

Cotesia typhae is a haplodiploid species. Females are produced from fertilized eggs and are diploid, whereas males are produced from unfertilized eggs and are haploid. We combined two cross schemes to produce F2 and F8 recombinant individuals (Figure 1). The production of F2 recombinant individuals was used to build a genetic map. This dataset was completed by F8 recombinant individuals to increase the number of recombination events for QTL detection. Three generations of sib-mating were first

realized for both Kobodo and Makindu parental lines. These sib-mating crosses following numerous generations (around 50 for Kobodo and 20 for Makindu) of rearing in small populations led to highly inbred parental lines. Virgin males and females from the parental strains were then crossed to generate F1 individuals. Part of the F1 females were isolated and kept virgin (Figure 1A). All other F1 individuals were mixed in a single population, which was maintained to reach F8 generation (Figure 1B). At each generation, all adults were stored in the same cage to allow random mating and randomly chosen females were used to produce the next generation to avoid overlapping between generations. Several F7 females were isolated and kept virgin. Virgin females from F1 or F7 generation were allowed to oviposit to produce respectively F2 and F8 recombinant males. For F2 males, each recombinant male was backcrossed with a single female from Kobodo or Makindu parental strains. For F8 males, only the Makindu strain was used, because a dominance of Kobodo alleles on these traits was susceptible to mask phenotypic variation as indicated by a previous study (Benoist et al., 2017). Since no meiosis occurs in males and parental strains are highly inbred, the offsprings from one recombinant male and one female from a parental line contain recombinant females with almost the same recombinant genotype at all loci (they were considered as genetically identical). Each female progeny was called a clonal sibship – CSS in the following text (Pannebakker et al., 2011). In total, 181 CSS were produced: 45 from F2 Kobodo backcross, 47 from F2 Makindu backcross, and 89 from F8 Makindu backcross. This experimental design allows replicated phenotyping for each recombinant genotype and thus the measurement of proportion traits on one CSS as well as parasitism success (see below), which requires several individuals to be assessed for accurate estimation, because of its variability.

2.3 Phenotyping

184 Between 15 and 20 females were tested for each CSS. In order to allow mating, sibling females were
185 left at least one day with their brothers in the same cage before the experiments (mating was not
186 controlled). The host caterpillars were placed individually under a 2cm diameter plastic top with one
187 female until the ovipositor insertion was observed. A female that refused to oviposit within 3 minutes
188 was discarded. Very few females were removed as host acceptance is very high (>90%) for the two
189 parasitoid strains (Benoist et al., 2020). After exposure, parasitized host caterpillars were kept in
190 rearing conditions until observation of either the formation of parasitoid cocoon mass (following the
191 emergence of the parasitoid larvae from the host), the death of the host without parasitoid emergence
192 or the formation of host pupa. After parasitoid emergence, each cocoon mass was placed in a tube to
193 obtain the adults. For each CSS, the parasitism success was calculated as the proportion of parasitized
194 hosts from which parasitoid larvae emerged. The two other traits were estimated, taking into account
195 successful parasitism only. The mean offspring number was calculated as the mean number of
196 parasitoid larvae that emerged from the host. The sex ratio for each CSS was determined by counting
197 the number of adult females, on one side, and the number of adult males, on the other, emerging from
198 each successful parasitism. To avoid bias, progenies of unmated females identified from their only
199 male content were not taken into account because the offspring number varies significantly between
200 mated and unmated females (Benoist et al., 2017). All phenotypic data are presented in the
201 supplementary Table S1. The broad-sense heritability in the CSS populations (H^2) could not be
202 estimated for parasitism success due to the lack of repetitions: only one measure is obtained for each
203 CSS based on the number of success and failure of the parasitism. By contrast, each success led to the
204 estimation of an offspring number and of a number of males and females. For these two traits, a
205 generalized linear model (GLM) was built, taking into account CSS and backcross type to explain the
206 trait variation. We used the “quasi-poisson” error family for the offspring number and the “quasi-

binomial" error family for the sex ratio. Genetic variance was estimated from the sum of squares associated with the CSS and it was divided by the total sum of squares to approach H^2.

2.4 RAD-Sequencing

A RAD-seq approach was performed in order to obtain genetic markers widely distributed in the genome. For each CSS, DNA was extracted from a pool of females using the NucleoSpin Tissue kit (Macherey-Nagel) and a RNase treatment was performed (Roche). All CSS DNA samples were digested using *PstI* restriction endonuclease. The expected number of restriction sites was approximately 18,500 based on *C. sesamiae kitale* genome analysis (Gauthier et al., 2020). A P1 adapter containing an Illumina adapter sequence, a 10 bp barcode (to identify CSS) and a sticky-end extremity, corresponding to the *PstI* site, was ligated to the *PstI*-digested fragments. F2 and F8 samples were pooled separately and adapter-ligated fragments were randomly sheared to obtain an average size of 600 bp. The two Illumina libraries were prepared by ligating a P2 adapter with a divergent end to DNA fragments to ensure that only fragments with P1 and P2 adapters would be fully amplified. These libraries were then amplified by PCR with P1 and P2 primers and paired-end sequenced on an Illumina NextSeq 500 instrument. The read length was 75 bp (including 10 bp of the barcode and 6 bp of the restriction site).

2.5 Identification of RAD locus and genotypes with Stacks

RAD-seq reads were trimmed to remove adaptors with Cutadapt v1.9.1 (Martin, 2011). RAD loci and the associated genotypes were determined using the Stacks v1.48 software package (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Rochette & Catchen, 2017). The trimmed reads were de-multiplexed and the barcodes were removed from reads using process radtags, discarding reads

with an uncalled base and/or low-quality score. F2 and F8 reads were treated separately. For each sample, reads were grouped in "stacks" to build loci using two approaches: *de novo* (ustacks) and reference-based (pstacks, in this case a reference genome is used to build loci). This double approach maximized the number of detected markers as the available reference genome is a draft version and is from a different species. Prior to *de novo* analyses, PCR duplicates were identified based on sequence identity and removed using home-made software. Preliminary tests were performed to optimize Stacks parameters: minimum depth coverage (-m) between 3 and 5, maximum distance between stacks (-M) between 2 and 3 and the number of mismatches allowed to build the catalog (-n) between 0 and 2 were tested. The selected parameters (m=3, M=2, n=2) were those that maximized the number of loci and minimized the variance between samples. For reference-based analyses, read pairs were mapped to the closely related *C. sesamiae kitale* genome (Gauthier et al., 2020) using BWA v0.7.17 (Li et al., 2009), and PCR duplicates sharing strict identical coordinates were removed with Samtools v1.9. The error rate upper bound was fixed at 0.01 for F2 (used to define the locus catalog, see below) and at 0.1 for F8. The deleveraging algorithm was used for *de novo* analyses. For all other parameters default settings were used. A *de novo* and a reference-based catalogue were generated from all F2 stacks (choosing population option) using the cstacks program with default parameters. Both F2 and F8 stacks were then matched to these catalogs to infer genotypes using the sstacks program with default parameters. We thus obtained two sets of genotypes for each CSS (F2 and F8): one from *de novo* and one from reference-based analysis.

2.6 Locus selection and genotype correction

The selection of markers from available loci was performed on F2 data. *De novo* and reference-based data were treated separately. Loci respecting the following conditions were selected: (i) diverging at 1

253 to 3 nucleotide sites between Kobodo and Makindu strains; (ii) showing invariance within each
254 parental line; (iii) found in at least 70 CSS; (iv) showing no segregation distortion. For all retained
255 markers, some genotypes of F2 and F8 CSS could be corrected by taking advantage of the backcross
256 design. For each backcross type, only two genotypes are expected among CSS: homozygous for
257 parental allele or heterozygous. For a CSS from the Makindu backcross, if a genotype was inferred as
258 homozygous for Kobodo alleles it was corrected for heterozygosity, assuming that the second allele
259 was not detected because of low coverage. The same procedure was applied for CSS from the Kobodo
260 backcross. After locus selection and genotype correction, *de novo* and reference-based data were
261 merged as follows. Genotypes were compared for the markers shared between both analyses. If
262 inferred genotypes differed between *de novo* and reference-based analyses, the heterozygous ones
263 were retained: we assumed that it was more likely to miss a second allele than to detect a false positive
264 one.

265 Only four markers showed segregation distortion. In total, 102,446 genotypes were infered (181
266 samples x 566 markers) and fewer than 1% were corrected based on *in silico* analysis. The comparison
267 of *de novo* and reference-based approaches led to the detection of 400 genotype discrepancies
268 between the two approaches, which contributed to the improvement of the genotypic data.

269

270 **2.7 Linkage map construction and correction**

271 The construction of the linkage map was performed with CarthaGene v1.3 (de Givry, Bouchez,
272 Chabrier, Milan, & Schiex, 2005) and the F2 data. The *LOD* and genetic distance thresholds used for
273 linkage group identification were respectively 7 and 0.5 Morgan. Marker ordering was done with the
274 default set of algorithms (Defalgo option). For each CSS, genotypes were organized following the
275 linkage map order to visualize recombination events and to correct some genotypes. When crossovers

were detected on both sides of a marker, the corresponding genotype was questioned. In this case, genotypes were encoded as missing values when coverage was below 10 (159 individual marker genotypes) or manually sequenced (86 individual marker genotypes). The final dataset comprised 3143 missing genotypes (3% of the data set). A new map was then generated from corrected data. When several markers were at null genetic distance on F2 and F8 data, the marker with more data was retained.

2.8 QTL analysis

All QTL analysis were performed with R Software (R Core Team, 2018). For each CSS, the probability of the genotypic states (KK, MM and KM for homozygous Kobodo genotypes, homozygous Makindu genotypes and heterozygotes respectively) at every cM map position were estimated using the package R/qtl v1.44-9 (Broman & Sen, 2009). At each position, additive and dominance indices were determined with the following formulas: *additive index* = $2P_{KK} + P_{KM}$; *dominance index* = P_{KM}, with P_{KK} and P_{KM} corresponding to the probability that the genotype was homozygous for Kobodo alleles or heterozygous.

To identify QTL for each trait, multiple regressions using GLM were performed. Due to the cross scheme, data could be classified into 3 backcross types which may influence their genotype: F2 backcross with Makindu strain, F2 backcross with Kobodo strain or F8 backcross with Makindu strain. As a consequence, phenotypic variation was analyzed in several steps. In a first step, a GLM was built for each trait, taking into account only the backcross type. They were based on the “quasi-binomial” error family for parasitism success and sex ratio and on the “gaussian” error family for the mean number of offsprings.

In a second step, all positions were scanned with a full GLM using residuals of the first model as response variable and additive and dominance indices of the given position as explanatory variables. Position *LOD* scores ($LOD_{Position}$) were calculated using the formula: $LOD_{Position} = n/2 \times log10 (RSS_{null}/RSS_{full})$, where *n* is the sample size, RSS_{full} is the residual sum of squares of the full model with additive and dominance indices, and RSS_{null} is the residual sum of squares of the null model (also based on residuals but without explanatory variables). The significant *LOD* score threshold was estimated by performing 2000 permutations between phenotypes and genotypes within cross types and taking the 5 % cut-off of the maximum *LOD* scores obtained as significant threshold value (Churchill & Doerge, 1994) The thresholds obtained were 3.06, 3.03, 3.01 for offspring number, parasitism success and sex ratio respectively.

In the next step, additive and dominance indices at the position with the highest significant *LOD* score were included to build a new model comprising this fixed position, and the genome was rescanned for an additional QTL. The process was repeated until no more significant position was detected. Such a process is especially powerful to detect genetically linked QTLs and QTLs in epistatic relationship. To test for interactions between QTLs, interaction terms for all QTL pairs were added to the model, including all QTLs detected. Each interaction term was then tested separately by an analysis of deviance comparing the model with and without interaction terms. The interaction was declared significant if the p-value was inferior to 0.05. For each trait, the final model that comprises all fixed QTLs and their significant pairwise interactions is called “the complete model” in the following text. In these models F8 backcross was the reference to calculate the intercept.

Additive and dominance coefficients are the estimated coefficients in these complete models. Additive effect corresponds to the variation associated with one Kobodo allele. Dominance effect corresponds to the variation resulting from the interaction between the parental alleles at one locus. Due to the “quasi-binomial” error family used for parasitism success models, estimated coefficients for this trait

are given in logits scale. These coefficients were transformed using the inverse logit transformation to estimate the parasitism success variation associated with additive and dominance effect (Crawley, 2012). The percentage of total phenotypic variance explained by each QTL was determined from the complete models with the formula SS_{QTL}/TSS, where SS_{QTL} is the sum of squares associated to QTL and TSS the total sum of squares. The confidence interval of each QTL position corresponded to all positions around the QTL for which *LOD* score was above *LOD* score max – 2.

2.9 Candidate gene identification

Gene identification was performed by genomic comparison with the available genome of *C. congregata*. We first assembled RADseq data from all samples with SPAdes v3.11.1 (Nurk et al., 2013). All markers included in the QTL confidence intervals were mapped on the scaffolds obtained using blastn v2.6.0+ (Boratyn et al., 2012). Scaffolds containing markers were then mapped to the annotated genome of *C. congregata* using BLAST to identify genes in QTLs. For this genome, an automatic annotation was performed and refined with manual inspection for some gene families (Bracovirus, Venom, Immunity, Detoxification and Chemodetection. Gauthier et al., 2020).

2.10 Gene Ontology enrichment analysis

Gene Ontology (GO) enrichment was tested using the R package topGO (Alexa, Rahnenfuhrer, & Lengauer, 2006). Annotations were those derived from blast2GO analyses of the *Cotesia congregata* genome performed by Gauthier et al. (2020). Two set of genes were built for each trait based of the gene content of QTL confidence interval. We compared the results of 4 algorithms implemented in topGO, namely 'classic', 'elim', 'weight' and 'weight01' that integrate the hierarchical structure of Gene

Ontology in different ways. Retaining only nodes larger than 5 we computed Fisher tests for the three categories of 'Biological Process', 'Molecular Function', and 'Cellular Compartment'.

3 Results

3.1 Descriptive Statistics

The two parental strains showed a substantial difference for parasitism success and offspring number, but they did not differ for sex ratio (Table 1). The Kobodo strain was more efficient with a higher parasitism success and a higher offspring number. The parasitism success of the clonal sibships (CSS) from the Kobodo backcross was equivalent to those of the Kobodo parental strain. By contrast, the parasitism success of the CSS from Makindu backcrosses (F2 and F8) was higher than the Makindu parental strain, although below the Kobodo. The offspring number of CSS from the Kobodo backcross was far higher than the Kobodo parental strain and the offspring number of the CSS from the Makindu backcrosses were close to those of the Kobodo parental strain. Data showed a higly significant correlation (Spearman's $\rho = 0.505$, p-value = 5.932×10^{-11}, calculated with all CSS) between the parasitism success and offspring number, suggesting a relationship between both traits. The broad-sense heritability was 29.82 % for offspring number and 14.46% for sex ratio.

3.2 Marker selection and genotype identification

The number of loci identified by *Stacks* varied considerably between the two approaches: 119,176 and 33,906 for *de novo* and reference-based, respectively (Table 2). However, the number of remaining loci after the first filter step was equivalent between the two approaches, around 30,000, which was similar to the 37,000 expected loci (based on the genome of the sister species *C. sesamiae*). This

suggests that a large proportion of loci are split into different stacks with the *de novo* approach. Among these 30,000 only a small fraction was polymorphic, suggesting that the two strains are genetically close. The majority of the identified markers were shared between the two approaches, but more than 20% (123/566) were specific to one of them, which highlighted their complementarity. The median coverage was above 30, which is satifying for reliable genotype inference.

3.3 Linkage map

The final linkage map generated from F2 data after removing redundant loci is presented in Figure 2 and in supplementary Table S2. It includes ten linkage groups, which is consistent with the number of chromosomes identifed in *C. typhae* (Christophe Bressac, personal communication), indicating that the map is saturated. In addition, each linkage group matched to one chromosome from *C. congregata* (Gauthier et al., 2020). The linkage map comprises 465 markers spanning a total length of 1300 cM with an average distance of 2.7 cM between two consecutive markers.

3.4 QTL mapping for parasitism success and offspring number

Three QTLs were identified for parasitism success (PS) and two for offsping number (ON), localized on linkage groups 4 and 6 and on the linkage groups 4 and 5, respectively (Figure 3). No QTL was detected for sex ratio. One peak for parasitism success was not retained (linkage group 1, two QTL scans) because it was observed in the first scan only and was located in a region with low marker density. The QTL observed on the linkage group 5 was common to the two phenotypic traits PS and ON. The length of the QTL confidence intervals varied from 9 to 38 cM (Table 3). For both PS and ON, the pairwise interactions between observed QTLs were not significant. An overdominance effect was observed for

all ON QTLs, whereas QTL1-PS and QTL3-PS were associated with respectively underdominance and partial dominance of Makindu alleles. The QTL2-PS was strictly additive. The percentage of phenotypic variance explained by each QTL varied between 4 and 14 %. The total phenotypic variance explained reached 27.7 % for parasitism success and 24.5 % for offspring number.

3.5 Identification of genes within QTL intervals

To identify candidate genes, markers in QTL intervals were mapped on the annotated genome of *C. congregata*. Cumulative length of all QTL intervals represented around 11.7 Mb (Table 4), which corresponds to ~5.9 % of the genome length and comprises 435 genes (Supplementary Table S3) with a putative function. This list of genes was inspected based on literature data. We focused on studies (mainly genomic and transcriptomic studies) that identified genes potentially influencing reproductive success in other parasitoid species. From these studies, 67 genes of interest were identified and are listed in Table 5. Among them we found Bracovirus genes involved in the production of viral particles and other bracovirus genes similar to genes involved in the inactivation of the host immune system in other host parasitoid models (Bézier, Annaheim, et al., 2009; Burke, Walden, Whitfield, Robertson, & Strand, 2014; Chevignon et al., 2015, 2014; Falabella et al., 2007; Gauthier et al., 2020; Pruijssers & Strand, 2007; Thoetkiattikul, Beck, & Strand, 2005). Some genes producing venom components were also identified (Ali, Lim, & Kim, 2015; Arvidson et al., 2019; Burke & Strand, 2014; Colinet, Mathé-Hubert, Allemand, Gatti, & Poirié, 2013; Danneels, Rivers, & de Graaf, 2010; Moreau & Asgari, 2015; Sim & Wheeler, 2016). Finally, we reported genes that were highly expressed in ovaries of *C. congregata* (Gauthier et al., 2020) and *N. vitripennis* (Sim & Wheeler, 2016) or differentially expressed between resting and ovipositing females of *N. vitripennis* (Pannebakker, Trivedi, Blaxter, Watt, & Shuker, 2013).

3.6 Gene Ontology enrichment analysis

Gene Ontology enrichment results are provided in Table 6. The GO terms were ordered following the results obtained with weight01 algorithm and the GO detected as significantly enriched at the 1% level with the four algorithms are listed. This study revealed an enrichment in GO terms involved in fine regulation processes. The most significant term refers to Calpain function, an intracellular calcium dependent cysteine protease (GO:0004198). This protease family is known to modulate the activity of other proteins, a function required in signal processing (Friedrich & Bozóky, 2005). In *Drosophila melanogaster*, Vieira, Cardoso, & Araujo (2017) showed that CalpA, a member of the calpain family, was involved in regulating the timing of mitosis during embryonic development. This molecular function may also favor the proper development of *C. typhae* larvae in its host. Terms involved in tRNA modifications (GO:0008175 and GO:0002098) were also detected as highly significant. Such post-transcriptional modifications are well known as modulators of tRNA activity influencing translation speed and fidelity. Repressing tRNA methylation decreased growth rates in yeast Nachtergaele & He (2017). Interestingly, RNA modifications are also widely used by viruses to hijack host cell machinery.

4 Discussion

The QTL approach is used to identify genes differing between lines diverging for one or several phenotypic traits. It is not designed to detect all the genes involved in the phenotype but rather the ones that vary and thus are susceptible to evolve. Applied to fitness related traits, this approach allows us to identify the genetic components involved in an organism's adaptation. In parasitoids, the QTL study of reproductive success provides an opportunity to identify the key features of host adaptation and hence the genes submitted to the co-evolution processes. Beyond the interest for understanding

parasitism success and reproduction in parasitoid species, such a study may also be helpful in the selection process in a biological control perspective, for example through marker-assisted selection. The QTL strategy is mainly conducted on model organisms, especially species of agricultural value, to identify genetic components of complex traits and also to support selection processes (Darvasi, 1998; Tanksley, 1993; Xu & Crouch, 2008). It was therefore challenging to develop such an approach on a recently described species.

The first challenge was to build a dense and reliable linkage map from scratch. The linkage map obtained comprise 10 linkage groups as the number of chromosomes observed in *C. typhae* (Christophe Bressac, personal communication) each of them corresponding to one of the 10 chromosomes of *C. congregata* (data not shown, Belle et al., 2002; Gauthier et al., 2020). The length of the genetic map is 1300 cM with an average recombination rate of 7.8 cM/Mb, in line with estimates for other insects (Stapley, Feulner, Johnston, Santure, & Smadja, 2017). The marker density combined with this recombination rate (one marker for each 400 Kbp) is favorable for QTL detection and confirms the interest of using RAD-seq to generate markers. This genetic map will be useful for further studies, for instance for the ongoing assembly of *C. typhae* genome at the chromosome scale.

The second challenge was to detect QTL, to characterize their phenotypic effect and to localize their genomic position as precisely as possible to obtain candidate genes. Owing to the favorable density of the map, the main limitation of QTL detection was the number of progenies to be characterized (Lander & Botstein, 1989). We phenotyped 181 CSS, each of them being characterized for 20 sibling females (more than 4000 parasitisms performed and more than 2600 successful progenies counted). This phenotyping effort allowed us to detect four distinct QTLs even with an impact below 5% of the phenotypic variance.

No QTL was detected for sex ratio despite a marked broad sense heritability estimated. This negative result probably arises from a complex determinism of sex ratio by numerous genes of low effect. In a comparable study on *Nasonia vitripennis*, Pannebakker et al. (2011) estimated H^2 of 9.5% for sex ratio and detected one QTL but explaining only 0.16 % of phenotypic variance and 1.56 % of genetic variance. These authors argued that sex ratio was likely to result from a complex architecture with pleiotropic genes influencing other life history traits such as clutch size. Such trade-off between traits was observed in *C. typhae* for which progenies from unmated females (with only male offsprings) comprised a significantly higher number of offsprings than progenies from mated females (mixed progenies with males and females, Benoist et al., 2017). This difference may result from ovipositing behavior of the female but also from the higher survival rate of male larvae compared to females. In such case, primary sex ratio (i.e. sex ratio at oviposition) should be a better trait to approach genetic determinism as it is directly imputable to the mother behavior (Ueno & Tanaka, 1997). However, this index is difficult to estimate, particularly for endoparasitoid species.

We identified three QTLs for parasitism success and two QTLs for offspring number. Comparison of the phenotypic variation explained by these QTLs with the broad-sense heritability estimated suggests that our approach succeeded in explaining almost all the expected genetic effect. This conclusion should be mitigated by the known bias of QTL detection strategies that result in an overestimate of the QTL effect through selection bias (Broman, 2001). Understanding of the genetic architecture of the host-parasite interaction has been approached mainly through the study of host genetic variation. In plant-parasite relationships, a gene-for-gene model has been proposed where a single locus is involved in the resistance of the host and a single locus involved in the virulence of the parasite (Thompson & Burdon, 1992). Based on genetic crosses between virulent and avirulent strains of the parasitoid *Leptopilina boulardi* and tests of parasitism success against different hosts of *Drosophila,* the same gene-for-gene model was proposed by Dupas et al. (1998) and Dupas and Carton (1999) to explain the

outcome of the interaction between the two partners. Mochiah, Ngi-Song, Overholt, and Stouthamer (2002) conducted crosses between *Cotesia sesamiae* strains differing in their ability to parasitize *Buseola fusca*. They also tested the relative impact of maternal factors and larval ones through superparasitization experiments. They identified a higher number of segregating factors and showed that both maternal factors and larval components (for example surface proteins) were necessary to allow the complete development of the parasitoid and the emergence of adults. Other theoretical models of co-evolution have been proposed that also involve more than one locus with strong epistatic relationship between loci (Tellier & Brown, 2007). The number of loci we identified is consistent with such a model except we did not find evidence of interactions between QTLs. However, the number of QTLs has to be analyzed cautiously as one QTL is not synonymous with one gene but may comprise several genes. The observation of enrichment for some GO terms in the QTL intervals suggests that several genes contribute to the variation in phenotype.

Besides the QTL detection, our approach allowed us estimating the additive and dominance effect for each QTL. Strong overdominance effects were observed for offspring number QTLs in agreement with results obtained on hybrid F1 (Benoist et al., 2017). Such effects reveal a positive interaction between alleles, either from a single gene or from different genes in the QTL interval through epistatic interactions or pseudo-superdominance. Under the pseudo-superdominance hypothesis, the overdominance effect is likely due to favorable alleles in the repulsion phase at tightly linked QTLs as observed in maize (Larièpe et al., 2012), suggesting that both Kobodo and Makindu may carry favorable alleles for offspring number. Parasitism success results also reveal the presence of favorable alleles in the Makindu strain, despite its overall poorer performance: negative values were estimated for dominance effect of QTL1-PS and additive effect of QTL3-PS. Taken together, these results indicate that a recombinant strain between Kobodo and Makindu, harboring all favorable alleles, may perform better than the parental strains and could be useful in biocontrol perspective.

504 The detection of a QTL shared between parasitism success and offspring number is consistent with the
505 high correlation detected between the two traits. In our experiment, the number of oviposition events
506 was fixed at one. The offspring number thus depends on (i) the number of eggs laid in the host by the
507 female and (ii) the parasitoid larval survival rate. Both larval mortality and parasitism success are
508 directly connected to parasitoid ability to inhibit the host immune system and are therefore naturally
509 correlated. In addition, some authors have even suggested that parasitoid females injecting more eggs
510 enhance the survival rate of their larvae through the saturation of the host immune system (Blumberg
511 & Luck, 1990; Kapranas, Tena, & Luck, 2012; Rosenheim & Hongkham, 1996). It is therefore not
512 surprising that the reproductive success is under the control of at least one common QTL between
513 parasitism success and offspring number.

514 In previous studies we found that the difference in offspring number observed between Kobodo and
515 Makindu females can be explained in part by the number of eggs injected into the host (Benoist et al.,
516 2017, 2020). Thus, we expected to find genes related to the oviposition behavior, especially in the QTL
517 specific to offspring number. However, very few genes influencing such behavior are known despite
518 the large number of studies on this topic. This is probably due to the complexity of behavioral traits
519 that involve a huge quantity of genes, complicating their study (Flint, Greenspan, & Kendler, 2010). A
520 refined annotation of genes in the QTL intervals is therefore required to identify candidate genes
521 related to oviposition behavior.

522 Genes belonging to the bracovirus were found in the common QTL and in the QTL3-PS. Within the
523 wasp genome the bracovirus is organized in two types of regions. The first corresponds to genes from
524 a nudiviral origin (nudiviral genes) involved in the bracovirus particle production. Numerous nudiviral
525 genes are clustered in one region, called the nudiviral cluster (Bézier, Annaheim, et al., 2009; Bézier,
526 Herbinière, et al., 2009) encoding in particular major capsid components (VP39, 38K, Wetterwald et
527 al., 2010). The second type corresponds to proviral segments used for the production of the viral circle

528 contained in viral particles and harboring virulence genes. In *C. congregata,* 35 proviral segments
529 organized in 9 proviral loci (PL) were identified. Approximately two-thirds of the proviral segments are
530 localized at the same genomic region, known as macrolocus, and comprising PL1 and PL2, while other
531 are dispersed in the wasp genome (Bézier et al., 2013). The adaptive role of the bracovirus in host
532 range evolution was widely documented in the *Cotesia* genus. For example, virulence of *C. sesamia*
533 against *Buseola fusca* was linked to allelic variation of the CrV1 bracovirus genes (Branca, Le Ru, Vavre,
534 Silvain, & Dupas, 2011; Gitau, Gundersen-Rindal, Pedroni, Mbugi, & Dupas, 2007). Furthermore,
535 several bracovirus genes were shown to be under positive selection in relation with host adaptation,
536 notably in *C. sesamiae* and *C. typhae* (Gauthier et al., 2018; Jancek et al., 2013).

537 In the QTL common to both traits we found proviral genes: BV20 genes and genes belonging to EP1-
538 like, PTP and Vank families for which some members are known as virulence factors of other parasitoid
539 species (Gueguen, Kalamarz, Ramroop, Uribe, & Govind, 2013; Kwon & Kim, 2008; Pruijssers & Strand,
540 2007). These genes are localized in PL5 and PL8 producing circles 1 and 26 respectively. These two
541 circles are of particular interest. It was shown in the *Cotesia congregata-Manduca sexta* system that
542 they could integrate into the host genome (Chevignon et al., 2018). They are produced in higher
543 numbers than others (Chevignon et al., 2014) and contain genes highly transcribed in the host. For
544 instance, EP1-like genes and Vank9 are among the most expressed in host haemocytes following
545 parasitism (Chevignon et al., 2014). Genes on these circles are thus good candidates to explain the
546 difference of parasitism success and offspring number between Kobodo and Makindu parasitoid
547 strains. Interestingly, no QTL was detected in the macrolocus region, which concentrates the majority
548 of proviral genes and no virulence genes known to be under positive selection were found in our QTL.

549 Many studies on virulence, with respect to bracovirus, focus on proviral genes but much less on
550 nudiviral genes encoding particle structural components and/or involved in particle production.
551 Interestingly, the QTL3-PS encompasses the whole nudiviral cluster. Among the genes in this QTL some

552 have a predicted function based on homology with baculovirus genes: *pif3* and *pif6* gene products may
553 play a role in virus entry into host cells, whereas VP39 and 38K likely produce major components of
554 nucleocapsids (Bézier, Annaheim, et al., 2009; Herniou, Olszewski, Cory, & O'Reilly, 2003; Wetterwald
555 et al., 2010) containing bracovirus DNA circles. Benoist et al., 2020 observed that the amount of viral
556 particles injected in the host did not explain the difference in parasitism success between Kobodo and
557 Makindu parasitoid strains, which make particle component production unlikely to be involved in the
558 difference between the two strains. The variation of parasitism success induced by this QTL could result
559 from a difference in particle infectivity between the two *C. typhae* strains, which in turn may result
560 from differences in *pif3* and/or *pif6* copies. It would be interesting therefore to compare how Kobodo
561 and Makindu bracovirus infect host cells.

562 In all QTLs, except in the QTL3-PS, we found genes associated to venom components, whose role in
563 host adaptation was widely studied (Cavigliasso et al., 2019). In polydnavirus-associated parasitoids,
564 polydnavirus is considered as the main virulence factor. However, it was shown in many species
565 harboring polydnavirus that venom is also required for successful parasitism and could synergize the
566 effect of polydnavirus (Asgari, 2012; Kitano, 1986; Moreau & Asgari, 2015; Tanaka, 1987). The presence
567 of venom genes in the QTLs suggests that their role in virulence might be significant and would
568 therefore make further investigations worthwhile.

569 The goal of this analysis was to identify candidate genes. In total 67 genes of interest were retained,
570 which is rather high for undertaking further studies of their individual implication in phenotypic
571 variation. Complementary approaches, such as comparative transcriptomics or genome wide
572 association study (GWAS) could enable us to select the most interesting candidate genes. Compared
573 to classic QTL approaches, GWAS allow to reach a higher resolution as soon as the sampling effort in
574 mixed or natural populations and the marker density are large enough (Hansson et al., 2018; Santure
575 & Garant, 2018). Focusing on QTL confidence interval, it may help to better target genes of interest.

Once the number of candidates genes is reduced, we will be able to assess their role using functional analyses available in this model, such as RNA interference; an approach which was shown to work efficiently to knock down targeted gene expression in Hymenoptera (Marco Antonio, Guidugli-Lazzarini, do Nascimento, Simões, & Hartfelder, 2008) and more specifically in parasitoid wasps (Burke, Thomas, Eum, & Strand, 2013; Colinet et al., 2014).

5 Conclusion

This work was devoted to the study of genetic components of the reproductive success of a parasitoid species. It pointed out four genomic regions involved in the variations of both parasitism success and offspring number, two traits directly connected to the fitness of individuals. It allowed the identification of a list of genes of interest, notably including bracovirus and venom genes. The detection of those well-known components of parasitoid virulence gives strong support to the strategy presented here. Those genes are particularly interesting in the topic of co-evolution because of their implication in host adaptation. The number of genes pointed out is quite large but clearly limited, their location being restricted by the QTL intervals to well-defined genomic regions. Population studies taking benefit of linkage disequilibrium at small genomic scale, or comparative transcriptomic studies, will allow to approach closer to the candidate genes in the future.

Acknowledgements

We thank Odile Giraudier and Sylvie Nortier for insect rearing at Gif, Julius Obonyo for field collection in Kenya, Matthieu Bodet, Florian Decourcelle, Maxime Villoing and Laurence Signon for their contribution to the crosses and phenotyping, Maud Silvain, Claude Thermes, Delphine Naquin and Yan

598 Jaszczyszyn from the I2BC sequencing facility for the RAD-sequencing, Florence Prunier for her
599 methodological advice, Damien Delafoy for his preliminary work on RAD-seq data, Cécile Courret for
600 her help in QTL analysis, Jérémy Gauthier for his advice on sequencing data analysis and his help with
601 the annotated genome of *C. congregata*, Lionel Saunois, Amandine Dubois and Virginie Héraudet for
602 maize production and Malcolm Eden for linguistic editing of the manuscript.

603

604 Data accessibility

605 RAD-Seq raw data sequences are archived at the NCBI SRA in the BioProject PRJNA622407. Genome
606 database (genomes and annotated genes) are available on the web site BIPAA (Bioinformatic
607 Platform for Agrosystem Arthropods) https://bipaa.genouest.org/is/parwaspdb/.

608

609 Author contributions

610 **Conceptualization**: R.B., S.D., F.M.; **Methodology**: R.B., C.C-D., L.M., A.L.R., L.K., F.M.; **Software**: R.B.,
611 C.C-D., F.M.; **Validation**: R.B, C.C-D., F.M.; **Formal analysis**: R.B., C.C-D., L.M., A.L.R., F.M.;
612 **Investigation**: R.B., C.C-D., C.C., R.J., P-A.C., E.V.D., L.K., F.M.; **Resources**: B.L.R, P-A.C, J-M.D., L.K.; **Data**
613 **curation**: R.B, C.C-D., F.M.; **Writing - original draft**: R.B., F.M.; **Writing - review & editing**: C.C-D., L.M.,
614 A.L.R., P-A.C., E.V.D., S.D., J-M.D., B.L.R, L.K.; **Supervision**: L.K., F.M.; **Project administration**: L.K.;
615 **Funding acquisition**: L.K.

616

617 Funding

618 This work was supported by the French National Research Agency (ABC Papogen project ANR-12-
619 ADAP-0001 and project Cotebio ANR-17-CE32-0015), and by the authors' operating grants from IRD,

CNRS, and icipe. R. Benoist is funded by the *Ecole doctorale 227 MNHN-UPMC Sciences de la Nature et de l'Homme: évolution et écologie*. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Additional information

All experimentations were realized under the juridical frame of a Material Transfer Agreement signed between IRD, icipe and CNRS (CNRS 072057/IRD 302227/00) and the authorization to import Cotesia in France delivered by the DRIAAF of Ile de France (IDF 2017-OI-26-032).

Supporting information

Table S1 Phenotypic data

Table S2 Detailed information of linkage map

Table S3 QTL gene content

633 **References**

634 Alexa, A., Rahnenfuhrer, J., & Lengauer, T. (2006). Improved scoring of functional groups from
635 gene expression data by decorrelating GO graph structure. *Bioinformatics*, *22*(13),
636 1600–1607. doi: 10.1093/bioinformatics/btl140

637 Ali, M. R., Lim, J., & Kim, Y. (2015). Transcriptome of a specialized extra-embryonic cell,
638 teratocyte, and its host immunosuppressive role revealed by *ex vivo* RNA interference:
639 Teratocyte transcriptome. *Insect Molecular Biology*, *24*(1), 13–28. doi:
640 10.1111/imb.12132

641 Arvidson, R., Kaiser, M., Lee, S. S., Urenda, J.-P., Dail, C., Mohammed, H., … Adams, M. E.
642 (2019). Parasitoid jewel wasp mounts multipronged neurochemical attack to hijack a
643 host brain. *Molecular & Cellular Proteomics*, *18*(1), 99–114. doi:
644 10.1074/mcp.RA118.000908

645 Asgari, S. (2012). Venoms from endoparasitoids. In *Parasitoid Viruses* (pp. 217–231). doi:
646 10.1016/B978-0-12-384858-1.00018-7

647 Asgari, S., & Rivers, D. B. (2011). Venom proteins from endoparasitoid wasps and their role in
648 host-parasite interactions. *Annual Review of Entomology*, *56*(1), 313–335. doi:
649 10.1146/annurev-ento-120709-144849

650 Beckage, N. E., & Drezen, J.-M. (Eds.). (2012). *Parasitoid viruses: symbionts and pathogens* (1st
651 ed). London ; New York: Elsevier/Academic Press.

652 Beckage, N. E., & Gelman, D. B. (2004). Wasp parasitoid disruption of host development:
653 implications for new biologically based strategies for insect control. *Annual Review of*
654 *Entomology*, *49*, 299–330. doi: 10.1146/annurev.ento.49.061802.123324

655 Belle, E., Beckage, N. E., Rousselet, J., Poirie, M., Lemeunier, F., & Drezen, J.-M. (2002).
656 Visualization of polydnavirus sequences in a parasitoid wasp chromosome. *Journal of*
657 *Virology*, *76*(11), 5793–5796. doi: 10.1128/JVI.76.11.5793-5796.2002
658 Benoist, R., Chantre, C., Capdevielle-Dulac, C., Bodet, M., Mougel, F., Calatayud, P. A., … Kaiser,
659 L. (2017). Relationship between oviposition, virulence gene expression and parasitism
660 success in *Cotesia typhae nov. sp.* parasitoid strains. *Genetica*. doi: 10.1007/s10709-
661 017-9987-5
662 Benoist, R., Paquet, S., Decourcelle, F., Guez, J., Jeannette, R., Calatayud, P.-A., … Kaiser, L.
663 (2020). Role of egg-laying behavior, virulence and local adaptation in a parasitoid's
664 chances of reproducing in a new host. *Journal of Insect Physiology*, *120*, 103987. doi:
665 10.1016/j.jinsphys.2019.103987
666 Bézier, A., Annaheim, M., Herbinière, J., Wetterwald, C., Gyapay, G., Bernard-Samain, S., …
667 Drezen, J.-M. (2009). Polydnaviruses of braconid wasps derive from an ancestral
668 nudivirus. *Science (New York, N.Y.)*, *323*(5916), 926–930. doi:
669 10.1126/science.1166788
670 Bézier, A., Herbinière, J., Lanzrein, B., & Drezen, J.-M. (2009). Polydnavirus hidden face: the
671 genes producing virus particles of parasitic wasps. *Journal of Invertebrate Pathology*,
672 *101*(3), 194–203. doi: 10.1016/j.jip.2009.04.006
673 Bézier, A., Louis, F., Jancek, S., Periquet, G., Theze, J., Gyapay, G., … Drezen, J.-M. (2013).
674 Functional endogenous viral elements in the genome of the parasitoid wasp *Cotesia*
675 *congregata*: insights into the evolutionary dynamics of bracoviruses. *Philosophical*

676 *Transactions of the Royal Society B: Biological Sciences*, *368*(1626), 20130047–
677 20130047. doi: 10.1098/rstb.2013.0047
678 Blumberg, D., & Luck, R. F. (1990). Differences in the rates of superparasitism between two
679 strains of *Comperiella bifasciata* (Howard) (Hymenoptera: Encyrtidae) parasitizing
680 California red scale (Homoptera: Diaspididae): an adaptation to circumvent
681 encapsulation? *Annals of the Entomological Society of America*, *83*(3), 591–597. doi:
682 10.1093/aesa/83.3.591
683 Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., & Madden, T. L.
684 (2012). Domain enhanced lookup time accelerated BLAST. *Biology Direct*, *7*(1), 12. doi:
685 10.1186/1745-6150-7-12
686 Branca, A., Le Ru, B. P., Vavre, F., Silvain, J.-F., & Dupas, S. (2011). Intraspecific specialization
687 of the generalist parasitoid *Cotesia sesamiae* revealed by polyDNAvirus polymorphism
688 and associated with different *Wolbachia* infection. *Molecular Ecology*, *20*(5), 959–971.
689 doi: 10.1111/j.1365-294X.2010.04977.x
690 Broman, K. W. (2001). Review of statistical methods for QTL mapping in experimental crosses.
691 *Lab Animal*, *30*(7), 44–52.
692 Broman, K. W., & Sen, S. (2009). *A guide to QTL mapping with R/qtl*. Dordrecht: Springer.
693 Burke, G. R., & Strand, M. R. (2014). Systematic analysis of a wasp parasitism arsenal.
694 *Molecular Ecology*, *23*(4), 890–901. doi: 10.1111/mec.12648
695 Burke, G. R., Thomas, S. A., Eum, J. H., & Strand, M. R. (2013). Mutualistic polydnaviruses share
696 essential replication gene functions with pathogenic ancestors. *PLoS Pathogens*, *9*(5),
697 e1003348. doi: 10.1371/journal.ppat.1003348

698 Burke, G. R., Walden, K. K. O., Whitfield, J. B., Robertson, H. M., & Strand, M. R. (2014).
699 Widespread genome reorganization of an obligate virus mutualist. *PLOS Genetics*,
700 *10*(9), e1004660. doi: 10.1371/journal.pgen.1004660
701 Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an
702 analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. doi:
703 10.1111/mec.12354
704 Cavigliasso, Mathé-Hubert, Kremmer, Rebuf, Gatti, Malausa, … Poirié. (2019). Rapid and
705 differential evolution of the venom composition of a parasitoid wasp depending on the
706 host strain. *Toxins*, *11*(11), 629. doi: 10.3390/toxins11110629
707 Charnov, E. L., & Skinner, S. W. (1985). Complementary approaches to the understanding of
708 parasitoid oviposition decisions. *Environmental Entomology*, *14*(4), 383–391. doi:
709 10.1093/ee/14.4.383
710 Chassain, C., & Bouletreau, M. (1987). Genetic variability in the egg-laying behaviour of
711 *Trichogramma maidis*. *Entomophaga*, *32*(2), 149–157. doi: 10.1007/BF02373125
712 Chevignon, G., Cambier, S., Da Silva, C., Poulain, J., Drezen, J.-M., Huguet, E., & Moreau, S. J.
713 M. (2015). Transcriptomic response of *Manduca sexta* immune tissues to
714 parasitization by the bracovirus associated wasp *Cotesia congregata*. *Insect*
715 *Biochemistry and Molecular Biology*, *62*, 86–99. doi: 10.1016/j.ibmb.2014.12.008
716 Chevignon, G., Periquet, G., Gyapay, G., Vega-Czarny, N., Musset, K., Drezen, J.-M., & Huguet,
717 E. (2018). *Cotesia congregata* bracovirus circles encoding *PTP* and *Ankyrin* genes
718 integrate into the DNA of parasitized *Manduca sexta* hemocytes. *Journal of Virology*,
719 *92*(15), e00438-18, /jvi/92/15/e00438-18.atom. doi: 10.1128/JVI.00438-18

720 Chevignon, G., Thézé, J., Cambier, S., Poulain, J., Silva, C. D., Bézier, A., … Huguet, E. (2014).
721 Functional annotation of *Cotesia congregata* bracovirus: identification of the viral
722 genes expressed in parasitized host immune tissues. *Journal of Virology*, JVI.00209-14.
723 doi: 10.1128/JVI.00209-14
724 Churchill, G. A., & Doerge, R. W. (1994). Empirical threshold values for quantitative trait
725 mapping. *Genetics*, *138*(3), 963–971.
726 Colinet, D., Kremmer, L., Lemauf, S., Rebuf, C., Gatti, J.-L., & Poirié, M. (2014). Development
727 of RNAi in a *Drosophila* endoparasitoid wasp and demonstration of its efficiency in
728 impairing venom protein production. *Journal of Insect Physiology*, *63*, 56–61. doi:
729 10.1016/j.jinsphys.2014.02.011
730 Colinet, D., Mathé-Hubert, H., Allemand, R., Gatti, J.-L., & Poirié, M. (2013). Variability of
731 venom components in immune suppressive parasitoid wasps: From a phylogenetic to
732 a population approach. *Journal of Insect Physiology*, *59*(2), 205–212. doi:
733 10.1016/j.jinsphys.2012.10.013
734 Crawley, M. J. (2012). *The R Book* (2nd Edition). Chichester, West Sussex, UK: Wiley-Blackwell.
735 Danneels, E. L., Rivers, D. B., & de Graaf, D. C. (2010). Venom proteins of the parasitoid wasp
736 *Nasonia vitripennis*: Recent discovery of an untapped pharmacopee. *Toxins*, *2*(4), 494–
737 516. doi: 10.3390/toxins2040494
738 Darvasi, A. (1998). Experimental strategies for the genetic dissection of complex traits in
739 animal models. *Nature Genetics*, *18*(1), 19–24. doi: 10.1038/ng0198-19

740 de Givry, S., Bouchez, M., Chabrier, P., Milan, D., & Schiex, T. (2005). CARHTA GENE:
741 multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics*,
742 *21*(8), 1703–1704. doi: 10.1093/bioinformatics/bti222
743 Drezen, J.-M., Leobold, M., Bézier, A., Huguet, E., Volkoff, A.-N., & Herniou, E. (2017).
744 Endogenous viruses of parasitic wasps: Variations on a common theme. *Current*
745 *Opinion in Virology*, *25*, 41–48. doi: 10.1016/j.coviro.2017.07.002
746 Dubuffet, A., Colinet, D., Anselme, C., Dupas, S., Carton, Y., & Poirié, M. (2009). Variation of
747 *Leptopilina boulardi* success in *Drosophila* hosts: What is inside the black box? In
748 *Advances in Parasitology*: *Vol. 70*. *Advances in Parasitology* (pp. 147–188). doi:
749 10.1016/S0065-308X(09)70006-5
750 Dupas, S., & Carton, Y. (1999). Two non-linked genes for specific virulence of *Leptopilina*
751 *boulardi* against *Drosophila melanogaster* and *D. yakuba*. *Evolutionary Ecology*, *13*(2),
752 211–220. doi: 10.1023/A:1006691431658
753 Dupas, S., Frey, F., & Carton, Y. (1998). A single parasitoid segregating factor controls immune
754 suppression in *Drosophila*. *The Journal of Heredity*, *89*(4), 306–311.
755 Edson, K. M., Vinson, S. B., Stoltz, D. B., & Summers, M. D. (1981). Virus in a parasitoid wasp:
756 Suppression of the cellular immune response in the parasitoid's host. *Science*,
757 *211*(4482), 582–583. doi: 10.1126/science.7455695
758 Falabella, P., Varricchio, P., Provost, B., Espagne, E., Ferrarese, R., Grimaldi, A., … Pennacchio,
759 F. (2007). Characterization of the I B-like gene family in polydnaviruses associated with
760 wasps belonging to different Braconid subfamilies. *Journal of General Virology*, *88*(1),
761 92–104. doi: 10.1099/vir.0.82306-0

762 Flint, J., Greenspan, R. J., & Kendler, K. S. (2010). *How Genes Influence Behavior* (1 edition).
763 Oxford ; New York: Oxford University Press.
764 Friedrich, P., & Bozóky, Z. (2005). Digestive versus regulatory proteases: On calpain action in
765 vivo. *Biological Chemistry*, *386*(7). doi: 10.1515/BC.2005.071
766 Gauthier, J., Boulain, H., van Vugt, J. J. F. A., Baudry, L., Persyn, E., Aury, J.-M., … Drezen, J.-M.
767 (2020). *Chromosomal resolution reveals symbiotic virus colonization of parasitic wasp*
768 *genomes* [Preprint]. doi: 10.1101/2020.03.19.994459
769 Gauthier, J., Gayral, P., LeRu, B. P., Jancek, S., Dupas, S., Kaiser, L., … Herniou, E. (2018). Genetic
770 footprints of adaptive divergence in the bracovirus of *Cotesia sesamiae* identified by
771 targeted re-sequencing. *Molecular Ecology*. doi: 10.1111/mec.14574
772 Gitau, C. W., Gundersen-Rindal, D., Pedroni, M., Mbugi, P. J., & Dupas, S. (2007). Differential
773 expression of the CrV1 haemocyte inactivation-associated polydnavirus gene in the
774 African maize stem borer *Busseola fusca* (Fuller) parasitized by two biotypes of the
775 endoparasitoid *Cotesia sesamiae* (Cameron). *Journal of Insect Physiology*, *53*(7), 676–
776 684. doi: 10.1016/j.jinsphys.2007.04.008
777 Godfray, H. C. J. (1994). *Parasitoids: Behavioral and evolutionary ecology*. Princeton University
778 Press.
779 Gueguen, G., Kalamarz, M. E., Ramroop, J., Uribe, J., & Govind, S. (2013). Polydnaviral ankyrin
780 proteins aid parasitic wasp qurvival by coordinate and selective inhibition of
781 hematopoietic and immune NF-kappa B signaling in insect hosts. *PLoS Pathogens*, *9*(8).
782 doi: 10.1371/journal.ppat.1003580

783 Hansson, B., Sigeman, H., Stervander, M., Tarka, M., Ponnikas, S., Strandh, M., … Hasselquist,
784 D. (2018). Contrasting results from GWAS and QTL mapping on wing length in great
785 reed warblers. *Molecular Ecology Resources*, *18*(4), 867–876. doi: 10.1111/1755-
786 0998.12785
787 Heimpel, G. E., & de Boer, J. G. (2008). Sex determination in the Hymenoptera. *Annual Review*
788 *of Entomology*, *53*(1), 209–230. doi: 10.1146/annurev.ento.53.103106.093441
789 Henter, H. J. (1995). The potential for coevolution in a host-parasitoid system. ii. Genetic
790 variation within a population of wasps in the ability to parasitize an aphid host.
791 *Evolution*, *49*(3), 439–445. doi: 10.1111/j.1558-5646.1995.tb02276.x
792 Herniou, E., Huguet, E., Theze, J., Bézier, A., Periquet, G., & Drezen, J.-M. (2013). When
793 parasitic wasps hijacked viruses: Genomic and functional evolution of polydnaviruses.
794 *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1626). doi:
795 10.1098/rstb.2013.0051
796 Herniou, E., Olszewski, J. A., Cory, J. S., & O'Reilly, D. R. (2003). The genome sequence and
797 evolution of baculoviruses. *Annual Review of Entomology*, *48*(1), 211–234. doi:
798 10.1146/annurev.ento.48.091801.112756
799 Jancek, S., Bézier, A., Gayral, P., Paillusson, C., Kaiser, L., Dupas, S., … Herniou, E. (2013).
800 Adaptive selection on bracovirus genomes drives the specialization of *Cotesia*
801 parasitoid wasps. *PLoS ONE*, *8*(5). doi: 10.1371/journal.pone.0064432
802 Kaiser, L, Dupas, S., Branca, A., Herniou, E., Clarke, C. W., Capdevielle Dulac, C., … Le Ru, B. P.
803 (2017). The *Cotesia sesamiae* story: Insight into host-range evolution in a

804 Hymenoptera parasitoid and implication for its use in biological control programs.
805 *Genetica*. doi: 10.1007/s10709-017-9989-3
806 Kaiser, L, Fernandez-Triana, J., Capdevielle-Dulac, C., Chantre, C., Bodet, M., Kaoula, F., … Le
807 Ru, B. (2017). Systematics and biology of *Cotesia typhae* sp. n. (Hymenoptera,
808 Braconidae, Microgastrinae), a potential biological control agent against the noctuid
809 Mediterranean corn borer, *Sesamia nonagrioides*. *ZooKeys*, *682*, 105–136. doi:
810 10.3897/zookeys.682.13016
811 Kaiser, L, Le Ru, B. P., Kaoula, F., Paillusson, C., Capdevielle-Dulac, C., Obonyo, J. O., … Dupas,
812 S. (2015). Ongoing ecological speciation in *Cotesia sesamiae*, a biological control agent
813 of cereal stem borers. *Evolutionary Applications*, *8*(8), 807–820. doi:
814 10.1111/eva.12260
815 Kaiser, Laure, Couty, A., & Perez-Maluf, R. (2009). Dynamic use of fruit odours to locate host
816 larvae: Individual learning, physiological state and genetic variability as adaptive
817 mechanisms. In *Advances in Parasitology* (Vol. 70, pp. 67–95). doi: 10.1016/S0065-
818 308X(09)70003-X
819 Kapranas, A., Tena, A., & Luck, R. F. (2012). Dynamic virulence in a parasitoid wasp: The
820 influence of clutch size and sequential oviposition on egg encapsulation. *Animal*
821 *Behaviour*, *83*(3), 833–838. doi: 10.1016/j.anbehav.2012.01.004
822 Kitano, H. (1986). The role of *Apanteles glomeratus* venom in the defensive response of its
823 host, *Pieris rapae* crucivora. *Journal of Insect Physiology*, *32*(4), 369–375. doi:
824 10.1016/0022-1910(86)90050-8

825 Kwon, B., & Kim, Y. (2008). Transient expression of an EP1-like gene encoded in *Cotesia*
826 *plutellae* bracovirus suppresses the hemocyte population in the diamondback moth,
827 *Plutella xylostella*. *Developmental & Comparative Immunology*, *32*(8), 932–942. doi:
828 10.1016/j.dci.2008.01.005
829 Lander, E. S., & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits
830 using RFLP linkage maps. *Genetics*, *121*(1), 185–199.
831 Larièpe, A., Mangin, B., Jasson, S., Combes, V., Dumas, F., Jamin, P., … Moreau, L. (2012). The
832 genetic basis of heterosis: Multiparental quantitative trait loci mapping reveals
833 contrasted levels of apparent overdominance among traits of agronomical interest in
834 maize (*Zea mays* L.). *Genetics*, *190*(2), 795–811. doi: 10.1534/genetics.111.133447
835 Le Masurier, A. D. (1987). A comparative study of the relationship between host size and brood
836 size in *Apanteles* spp. (Hymenoptera: Braconidae). *Ecological Entomology*, *12*(4), 383–
837 393. doi: 10.1111/j.1365-2311.1987.tb01019.x
838 Le Masurier, A. D. (1991). Effect of host size on clutch size in *Cotesia glomerata*. *The Journal*
839 *of Animal Ecology*, *60*(1), 107. doi: 10.2307/5448
840 Legner, E. F. (1987). Inheritance of gregarious and solitary oviposition in *Muscidifurax*
841 *raptorellus* kogan and legner (Hymenoptera: Pteromalidae). *The Canadian*
842 *Entomologist*, *119*(9), 791–808. doi: 10.4039/Ent119791-9
843 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project
844 Data Processing Subgroup. (2009). The sequence alignment/map format and
845 SAMtools. *Bioinformatics*, *25*(16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Mackay, T. F. C., Stone, E. A., & Ayroles, J. F. (2009). The genetics of quantitative traits: Challenges and prospects. *Nature Reviews Genetics*, *10*(8), 565–577. doi: 10.1038/nrg2612

Marco Antonio, D. S., Guidugli-Lazzarini, K. R., do Nascimento, A. M., Simões, Z. L. P., & Hartfelder, K. (2008). RNAi-mediated silencing of vitellogenin gene function turns honeybee (*Apis mellifera*) workers into extremely precocious foragers. *Naturwissenschaften*, *95*(10), 953–961. doi: 10.1007/s00114-008-0413-9

Marti, D. (2003). Ovary development and polydnavirus morphogenesis in the parasitic wasp *Chelonus inanitus*. I. Ovary morphogenesis, amplification of viral DNA and ecdysteroid titres. *Journal of General Virology*, *84*(5), 1141–1150. doi: 10.1099/vir.0.18832-0

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. doi: 10.14806/ej.17.1.200

Mochiah, M. B., Ngi-Song, A. J., Overholt, W. A., & Stouthamer, R. (2002). Wolbachia infection in *Cotesia sesamiae* (Hymenoptera: Braconidae) Causes cytoplasmic incompatibility: Implications for biological control. *Biological Control*, *25*(1), 74–80.

Moreau, S., & Asgari, S. (2015). Venom proteins from parasitoid wasps and their biological functions. *Toxins*, *7*(7), 2385–2412. doi: 10.3390/toxins7072385

Mrinalini, & Werren, J. H. (2016). Parasitoid wasps and their venoms. In P. Gopalakrishnakone & A. Malhotra (Eds.), *Evolution of Venomous Animals and Their Toxins* (pp. 1–26). doi: 10.1007/978-94-007-6727-0_2-2

Nachtergaele, S., & He, C. (2017). The emerging biology of RNA post-transcriptional modifications. *RNA Biology*, *14*(2), 156–163. doi: 10.1080/15476286.2016.1267096

868 Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., … Pevzner, P.
869 A. (2013). Assembling genomes and mini-metagenomes from highly chimeric reads. In
870 M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), *Research in Computational Molecular*
871 *Biology* (pp. 158–170). Springer Berlin Heidelberg.
872 Orzack, S. H., & Gladstone, J. (1994). Quantitative genetics of sex ratio traits in the parasitic
873 wasp, *Nasonia vitripennis*. *Genetics*, *137*(1), 211–220.
874 Pannebakker, B. A., Watt, R., Knott, S. A., West, S. A., & Shuker, D. M. (2011). The quantitative
875 genetic basis of sex ratio variation in *Nasonia vitripennis*: A QTL study. *Journal of*
876 *Evolutionary Biology*, *24*(1), 12–22. doi: 10.1111/j.1420-9101.2010.02129.x
877 Pannebakker, Bart A., Trivedi, U., Blaxter, M. A., Watt, R., & Shuker, D. M. (2013). The
878 transcriptomic basis of oviposition behaviour in the parasitoid wasp *Nasonia*
879 *vitripennis*. *PLoS ONE*, *8*(7), e68608. doi: 10.1371/journal.pone.0068608
880 Pennacchio, F., & Strand, M. R. (2006). Evolution of developmental strategies in parasitic
881 Hymenoptera. *Annual Review of Entomology*, *51*(1), 233–258. doi:
882 10.1146/annurev.ento.51.110104.151029
883 Pichon, A., Bézier, A., Urbach, S., Aury, J.-M., Jouan, V., Ravallec, M., … Volkoff, A.-N. (2015).
884 Recurrent DNA virus domestication leading to different parasite virulence strategies.
885 *Science Advances*, *1*(10), e1501150–e1501150. doi: 10.1126/sciadv.1501150
886 Pruijssers, A. J., & Strand, M. R. (2007). PTP-H2 and PTP-H3 from *Microplitis demolitor*
887 bracovirus localize to focal adhesions and are antiphagocytic in insect immune cells.
888 *Journal of Virology*, *81*(3), 1209–1219. doi: 10.1128/JVI.02189-06

889 Quicke, D. L. J. (2014). *The braconid and ichneumonid parasitoid wasps: Biology, systematics,*
890 *evolution and ecology*. doi: 10.1002/9781118907085
891 R Core Team. (2018). R: A language and environment for statistical computing. *R Foundation*
892 *for Statistical Computing, Vienna, Austria. URL Https://Www.R-Project.Org/.* Retrieved
893 from https://www.R-project.org/
894 Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data
895 using Stacks. *Nature Protocols*, *12*(12), 2640–2659. doi: 10.1038/nprot.2017.123
896 Rosenheim, J. A., & Hongkham, D. (1996). Clutch size in an obligately siblicidal parasitoid wasp.
897 *Animal Behaviour*, *51*(4), 841–852. doi: 10.1006/anbe.1996.0088
898 Santure, A. W., & Garant, D. (2018). Wild GWAS-association mapping in natural populations.
899 *Molecular Ecology Resources*, *18*(4), 729–738. doi: 10.1111/1755-0998.12901
900 Sim, A. D., & Wheeler, D. (2016). The venom gland transcriptome of the parasitoid wasp
901 *Nasonia vitripennis* highlights the importance of novel genes in venom function. *BMC*
902 *Genomics*, *17*(1), 571. doi: 10.1186/s12864-016-2924-7
903 Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation
904 in recombination frequency and distribution across eukaryotes: Patterns and
905 processes. *Philosophical Transactions of the Royal Society of London. Series B,*
906 *Biological Sciences*, *372*(1736). doi: 10.1098/rstb.2016.0455
907 Strand, M. R. (2014). Teratocytes and their functions in parasitoids. *Current Opinion in Insect*
908 *Science*, *6*, 68–73. doi: 10.1016/j.cois.2014.09.005
909 Strand, M. R., & Burke, G. R. (2015). Polydnaviruses: From discovery to current insights.
910 *Virology*, *479–480*, 393–402. doi: 10.1016/j.virol.2015.01.018

911 Tanaka, T. (1987). Calyx and venom fluids of *Apanteles kariyai* (Hymenoptera: Braconidae) as
912 factors that prolong larval period of the host, *Pseudaletia separata* (Lepidoptera:
913 Noctuidae). *Annals of the Entomological Society of America*, *80*(4), 530–533. doi:
914 10.1093/aesa/80.4.530
915 Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, *27*, 205–233. doi:
916 10.1146/annurev.ge.27.120193.001225
917 Tellier, A., & Brown, J. K. M. (2007). Polymorphism in multilocus host–parasite coevolutionary
918 interactions. *Genetics*, *177*(3), 1777–1790. doi: 10.1534/genetics.107.074393
919 Thoetkiattikul, H., Beck, M. H., & Strand, M. R. (2005). Inhibitor kappaB-like proteins from a
920 polydnavirus inhibit NF-kappaB activation and suppress the insect immune response.
921 *Proceedings of the National Academy of Sciences of the United States of America*,
922 *102*(32), 11426–11431. doi: 10.1073/pnas.0505240102
923 Thompson, J. N. (2009). *The Coevolutionary Process*. University of Chicago Press.
924 Thompson, J. N., & Burdon, J. J. (1992). Gene-for-gene coevolution between plants and
925 parasites. *Nature*, *360*(6400), 121–125. doi: 10.1038/360121a0
926 Ueno, T., & Tanaka, T. (1997). Comparison between primary and secondary sex ratios in
927 parasitoid wasps using a method for observing chromosomes. *Entomologia*
928 *Experimentalis et Applicata*, *82*(1), 105–108. doi: 10.1046/j.1570-7458.1997.00119.x
929 Velthuis, B.-J., Yang, W., Van Opijnen, T., & Werren, J. H. (2005). Genetics of female mate
930 discrimination of heterospecific males in *Nasonia* (Hymenoptera, Pteromalidae).
931 *Animal Behaviour*, *69*(5), 1107–1120. doi: 10.1016/j.anbehav.2004.06.032

Vieira, V., Cardoso, M. A., & Araujo, H. (2017). Calpain A controls mitotic synchrony in the *Drosophila* blastoderm embryo. *Mechanisms of Development*, *144*, 141–149. doi: 10.1016/j.mod.2016.05.005

Volkoff, A.-N., Jouan, V., Urbach, S., Samain, S., Bergoin, M., Wincker, P., … Drezen, J.-M. (2010). Analysis of virion structural components reveals vestiges of the ancestral ichnovirus genome. *PLOS Pathogens*, *6*(5), e1000923. doi: 10.1371/journal.ppat.1000923

Waage, J. K., & Greathead, D. J. (Eds.). (1986). *Insect parasitoids*. London ; Orlando: Academic Press.

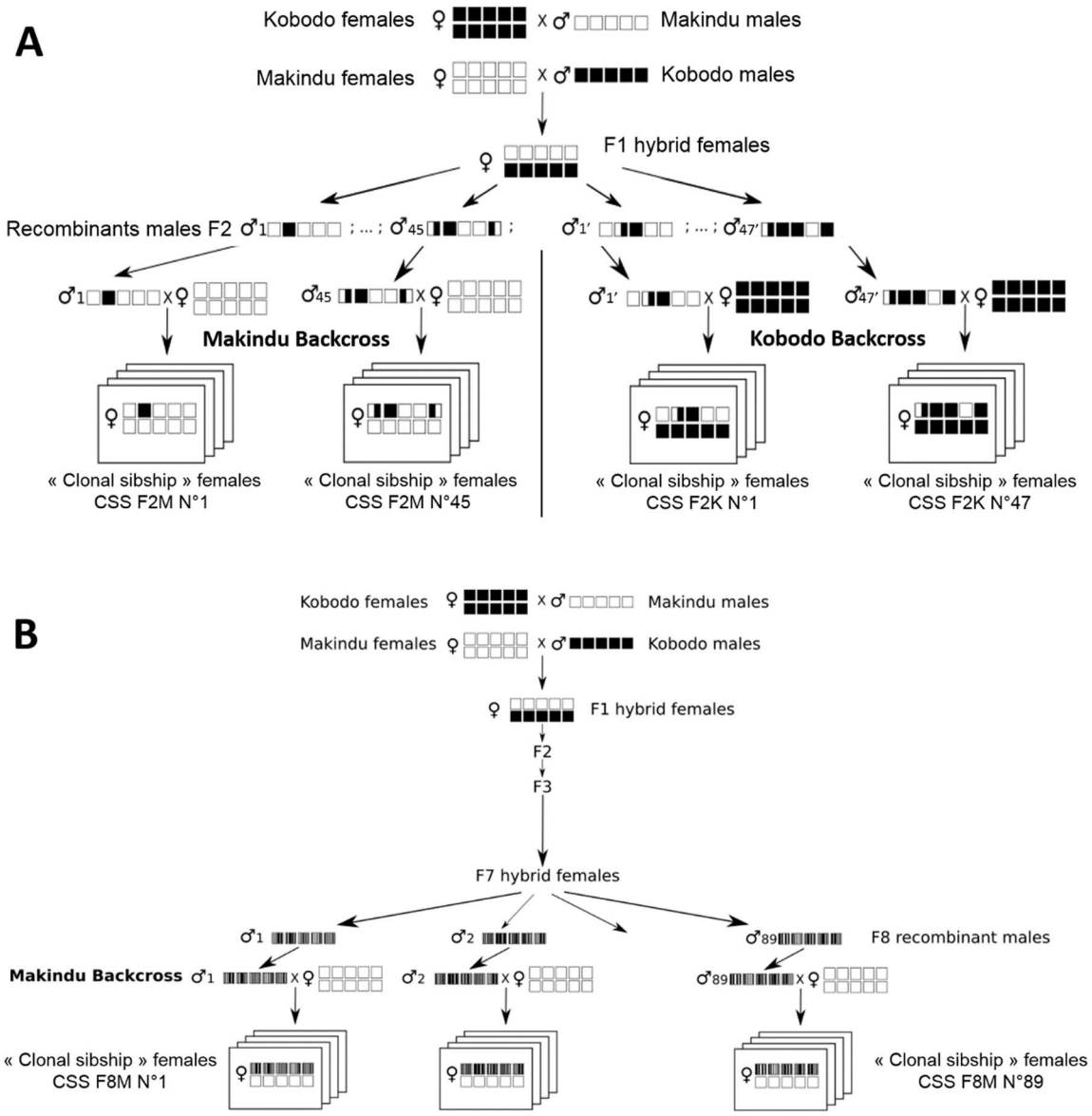
Waage, J. K., & Ming, N. S. (1984). The reproductive strategy of a parasitic wasp: I. Optimal progeny and sex allocation in *Trichogramma evanescens*. *The Journal of Animal Ecology*, *53*(2), 401. doi: 10.2307/4524

Wang, Z., Liu, Y., Shi, M., Huang, J., & Chen, X. (2019). Parasitoid wasps as effective biological control agents. *Journal of Integrative Agriculture*, *18*(4), 705–715. doi: 10.1016/S2095-3119(18)62078-7

Webb, B. A. (1998). Polydnavirus biology, genome structure, and evolution. In L. K. Miller & L. A. Ball (Eds.), *The Insect Viruses* (pp. 105–139). doi: 10.1007/978-1-4615-5341-0_5

Wetterwald, C., Roth, T., Kaeslin, M., Annaheim, M., Wespi, G., Heller, M., … Lanzrein, B. (2010). Identification of bracovirus particle proteins and analysis of their transcript levels at the stage of virion formation. *Journal of General Virology, 91*(10), 2610–2619. doi: 10.1099/vir.0.022699-0

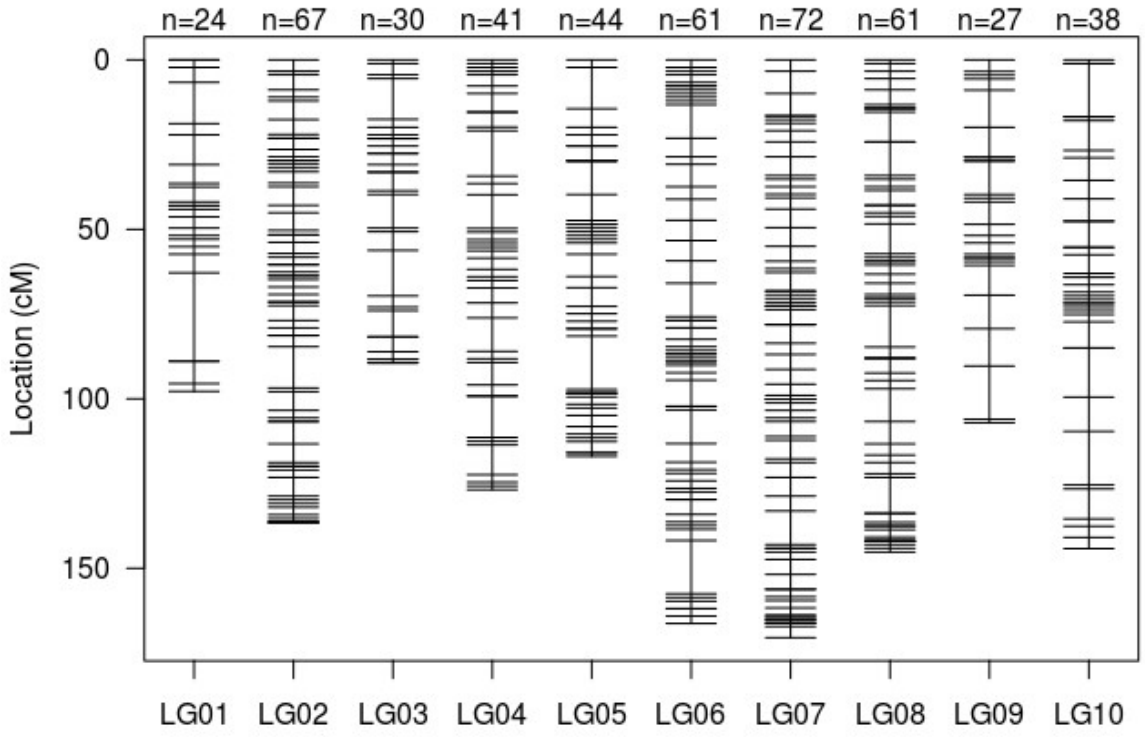
953 Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002).
954 Biological and biomedical implications of the co-evolution of pathogens and their
955 hosts. *Nature Genetics*, *32*(4), 569–577. doi: 10.1038/ng1202-569
956 Wyler, T., & Lanzrein, B. (2003). Ovary development and polydnavirus morphogenesis in the
957 parasitic wasp *Chelonus inanitus*. II. Ultrastructural analysis of calyx cell development,
958 virion formation and release. *Journal of General Virology*, *84*(5), 1151–1163. doi:
959 10.1099/vir.0.18830-0
960 Xu, Y., & Crouch, J. H. (2008). Marker-assisted selection in plant breeding: From publications
961 to practice. *Crop Science*, *48*(2), 391–407. doi: 10.2135/cropsci2007.04.0191
962



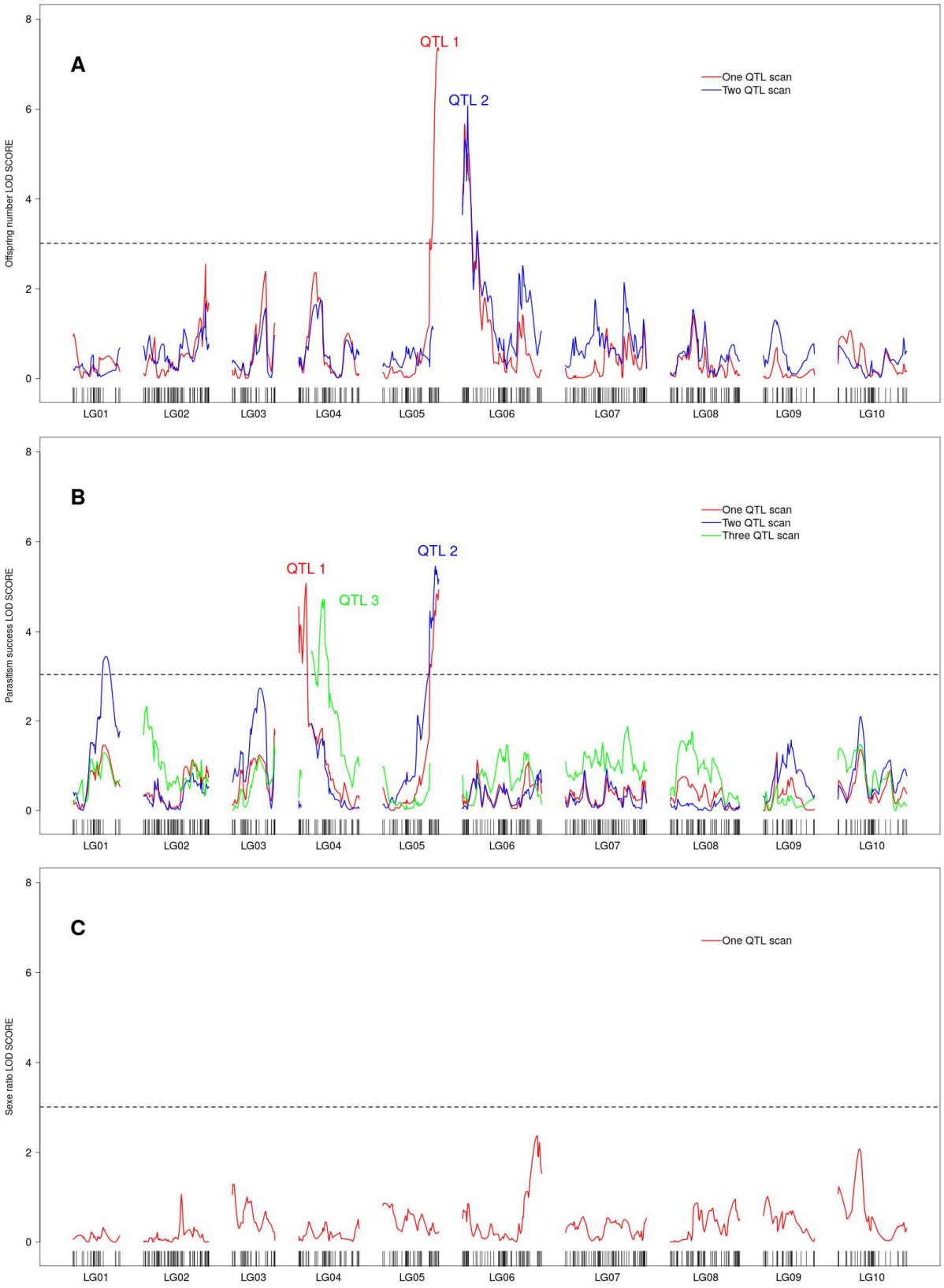
963 **Figure 1: Cross schemes used to generate clonal sibship, adapted from Pannebakker et al., 2011:** A:
964 F2 cross scheme; B: F8 cross scheme. *C. typhae* is a haplo-diploid species. A haploid genome set is
965 represented by 5 squares, with one set for males and two sets for females. Kobodo and Makindu
966 genetic contributions are represented in black and white respectively. Inbred parental strains are
967 considered homozygous. Kobodo and Makindu individuals from parental strains were crossed to
968 generate F1 individuals. A subset of F1 hybrid females were kept virgin to produce F2 recombinant
969 males (A). All other F1 individuals were mixed in a single population which was maintained until F8
970 generation (B). F2 (A) and F8 (B) recombinant males were then backcrossed with females from the
971 parental strains to produce clonal sibship (CSS) females (*i.e.* females considered as having identical
972 genotypes). Each rectangle inside one stack represents a female in a CSS. F8 males were backcrossed
973 only with Makindu females. In total, 181 CSS were produced: 45 from F2 Kobodo backcross, 47 from
974 F2 Makindu backcross, and 89 from F8 Makindu backcross. For each CSS, between 15 and 20 females

975 were used to measure the phenotypic traits of the CSS and all females were pooled for RAD-
976 sequencing.

977



978 **Figure 2: *Cotesia typhae* linkage map based on RAD genotyping of 92 F2 CSS**. The linkage map includes
979 465 markers, n=number of markers by linkage group (LG). Genetic distances are calculated using
980 Kosambi's map function.
981


A
QTL 1
QTL 2
One QTL scan
Two QTL scan
Offspring number LOD SCORE
B
QTL 1
QTL 2
QTL 3
One QTL scan
Two QTL scan
Three QTL scan
Parasitism success LOD SCORE
C
One QTL scan
Sexe ratio LOD SCORE
LG01
LG02
LG03
LG04
LG05
LG06
LG07
LG08
LG09
LG10

982

Figure 3: Result of *Cotesia typhae* genome scan for QTL detection of (A) offspring number, (B) parasitism success and (C) sex ratio. The 10 linkage groups (LG) and their markers are indicated on x-axis. *LOD* SCORES are calculated from generalized linear models (GLM). The black dotted line corresponds to the QTL threshold calculated by 2000 permutations: it reaches 3.06, 3.03, 3.01 for offspring number, parasitism success and sex ratio respectively. Successive scans were performed for each linkage group. In One QTL scan, we hypothesized that only one QTL occurred in the genome. The highest peak was then included in the two QTL scans to search for a second QTL. The process was repeated until no other significant peak was detected. Colored lines correspond to each QTL scan performed where a significant peak was detected.

Table 1: Phenotypes of the parental strains and the clonal sibships (CSS) from the different backcrosses. ± Standard deviation; n: number of female progenies analyzed; ( ) number of CSS; H^2=Broad-sense heritabiliy.

Data type	Parasitism Success		Offspring number		Sex ratio (% male)	
Parental Kobodo*	78.70%	n=108	71.07 ± 24.06	n=71	68.80 ± 20.22	n=71
Parental Makindu*	36.20%	n=80	36.48 ± 16.68	n=21	74.90 ± 16.50	n=21
CSS from Kobodo F2	81.57 ± 14.80 %	n=832 (47)	91.05 ± 17.54	n=491 (47)	39.09 ± 21.28	n=490 (47)
CSS from Makindu F2	70.96 ± 22.55 %	n=857 (45)	74.94 ± 24.77	n=440 (45)	35.89 ± 16.78	n=437 (45)
CSS from Makindu F8	64.00 ± 20.24 %	n=1737 (89)	66.83 ± 19.71	n=827 (88)	33.01 ± 20.00	n=801 (88)
			H^2=29.82%		H^2=14.46%	

* Data from Benoist et al. (2017)

998 **Table 2: Number of retained loci at each step of the loci selection and locus coverage.** The locus
999 coverage is given for the 566 markers used for the genetic map construction.

		Analysis	
		De novo	Reference based
Number of loci	Identified by Stacks	119,176	33,906
	Present in at least 70 samples	31,599	29,797
	Fixed in parental strains and polymorphic	541	507
	With at most 3 SNP	524	491
	Without segregation distortion	522	487
	De novo / Reference based specific	79	44
	Used for linkage map construction	566	
	Retained in final linkage map	465	
Marker coverage per sample	Mean	35.55	49.32
	Median	32	34

1000
1001

Table 3: Detected QTL position, confidence interval, coefficient estimates and percentage of the total phenotypic variance explained by each QTL detected. Additive effect corresponds to the effect of one Kobodo allele. Positive dominance effects indicate dominance of Kobodo alleles, whereas negative values reveal dominance of Makindu ones.

		Phenotypic trait				
		Offspring number		Parasitism success		
QTL name		**QTL1-ON**	**QTL2-ON**	**QTL1-PS**	**QTL2-PS**	**QTL3-PS**
Linkage group		LG05	LG06	LG04	LG05	LG04
LOD score value		6.71	5.91	5.08	5.44	4.68
Peak position		116 cM	11 cM	15 cM	110 cM	50 cM
Confidence interval		108-117 cM	2-17 cM	0-18 cM	98-117 cM	22-60 cM
Coefficient estimates	Additive effect	7.72	0.52	0.02	0.16	- 0.16
	Dominance effect	13.05	14.14	- 0.17	-*	0.08
% Total phenotypic variance explained	Additive component	8.29 %	2.55 %	4.68 %	11.15 %	3.50 %
	Dominance component	6.60 %	7.06 %	7.62 %	-*	0.75 %
	Total by QTL	14.89 %	9.61 %	12.30 %	11.15 %	4.25 %
	Total	24.5 %		27.7 %		

*****The percentage of the total phenotypic variance explained by the dominance component of the QTL2-PS was not calculated due to the lack of significance of this component in the QTL model.

Table 4: Summary of gene identification in quantitative trait loci (QTL) intervals based on the annotated genome of *C. congregata*. Overlapping QTLs (QTL1-ON and QTL2-PS) was considered as a single interval.

QTL name	*C. congregata* fragment length	Number of genes		
		Automatically annotated	Manually annotated	Total
QTL1-PS	~ 1.3 Mb	47	15	62
QTL3-PS	~ 3.7 Mb	105	37	142
QTL1-ON + QTL2-PS	~ 4.2 Mb	113	25	138
QTL2-ON	~ 2.5Mb	90	3	93

Table 5: Genes of interest in quantitative trait loci (QTL) intervals. Overlapping QTL (QTL1-ON and QTL2-PS) was considered as a single interval. Genes of interest were classified in four categories. Bracovirus Nudiviral cluster (genes of nudiviral origin involved in particles bracovirus production): Bracovirus virulence genes (genes packaged in the particles and expressed during parasitism) and Venom categories refer to parasitoid weapon arsenal. The "Other" category comprises genes selected because they are actively transcribed in ovaries of *C. congregata* (Gauthier et al., 2020) and *N. vitripennis* (Sim & Wheeler, 2016) or differentially expressed between resting (no contact with host) and ovipositing females of *N. vitripennis* (Bart A. Pannebakker et al., 2013).

QTL name	Genes of interest
QTL1-PS	**Venom:** alkaline phosphatase-like; disintegrin and metalloproteinase domain-containing protein 9 **Other**: btb poz domain-containing adapter for cul3-mediated degradation protein 3; 60s ribosomal protein l10; 60s ribosomal protein l23a
QTL3- PS	**Bracovirus nudiviral cluster**: 27b; 35a$_1$; 35a$_2$; 38K; GbNVorf19; HzNVorf106; HzNVorf9_1; HzNVorf9_2; HzNVorf94; Int$_1$; K425_438; K425 445; K425 456; K425 459; K425 461; P6.9$_1$; P6.9$_2$; pif3; pif6; PmNVorf87; PmV; ToNVorf29; ToNVorf54$_1$; ToNVorfF54$_2$; vp39 **Other** : Putative mediator of RNA polymerase II transcription subunit 12; 40s ribosomal protein s15aa
QTL1-ON + QTL2-PS	**Bracovirus virulence genes**: bv20.1.26.4; bv20.2.26.8; ep1.1.1.3; ep1.2.1.4; ep1.3.1.5; ptp a.26.6; ptp b.1.1; ptp d.1.11 pseudo; ptp delta.26.1; ptp epsilon.26.7 pseudo; ptp i.1.2; ptp k.1.6; ptp l.1.7; ptp m.1.8; ptp p.1.9; ptp q.1.10; vank 5.26.2; vank 6.26.3b; vank 9.26.5 **Venom**: calcium-independent phospholipase a2; adenosine deaminase-like protein; cysteine-rich with EGF-like domain protein 2; serine protease inhibitor; serpin B4-like; serpin B6-like **Other**: plasminogen activator inhibitor 1 RNA-binding protein; BTB/POZ domain-containing protein 7; sorting-nexin 24-like; sorting-nexin 6; zinc finger and BTB domain-containing protein 41-like; 60S ribosomal protein l13; 60S ribosomal protein l18a
QTL2-ON	**Venom**: cysteine-rich secretory protein 2-like **Other**: guanine nucleotide-binding protein subunit beta-like protein; 60S ribosomal protein l5

Table 6: Summary of GO enrichment tests in quantitative trait loci (QTL) intervals based on the annotated genome of *C. congregata*. The list comprises GO detected as significantly enriched at the 1% level with 4 different algorithms implemented in topGO package: classic, elim, weight and weight01. P-values indicated are those obtained with weight01 package.

Trait	GO,ID	Term	Observed number of genes	Expected number of genes	p-value
Parasitism success		**Molecular Function**			
	GO:0004198	calcium-dependent cysteine-type endopeptidase activity	3	0.21	0.00069
	GO:0008175	tRNA methyltransferase activity	3	0.29	0.00227
	GO:0004674	protein serine/threonine kinase activity	11	4.17	0.0027
		Biological Process			
	GO:0002098	tRNA wobble uridine modification	4	0.25	0.000043
	GO:0000902	cell morphogenesis	3	0.34	0.0035
	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	3	0.34	0.0035
	GO:0030488	tRNA methylation	3	0.34	0.0035
	GO:0120036	plasma membrane bounded cell projection organization	3	0.42	0.0099
Offspring Number		**Molecular Function**			
	GO:0004198	calcium-dependent cysteine-type endopeptidase activity	2	0.16	0.0093
	GO:0030246	carbohydrate binding	4	0.85	0.0095
		Biological Process			
	GO:0042176	regulation of protein catabolic process	3	0.23	0.0011
	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	3	0.26	0.0017
	GO:0120036	plasma membrane bounded cell projection organization	3	0.33	0.006