The for gene as one of the drivers of foraging variations in a parasitic wasp

aurore gallot[1], Emmanuel Desouhant[2], Vincent Lhuillier[1], david lepetit[1], Adil El-filali[3], Laurence Mouton[2], Cristina Vieira[2], and Isabelle Amat[4]

[1]Université Claude Bernard Lyon 1
[2]Université de Lyon
[3]Université Lyon 1, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, Villeurbanne, France
[4]Biometrie et biologie evolutive

September 1, 2022

Abstract

Foraging behaviours encompass strategies to locate resources and to exploit them. In many taxa these behaviours are controlled by a major gene called for, but mechanisms vary between species. In the parasitoid wasp Venturia canescens, sexual and asexual populations coexist in sympatry but differ in their foraging behaviours. Here we explored the molecular bases underpinning this divergence in foraging behaviours by testing two mutually non-exclusive hypotheses: firstly the divergence in the for gene results in difference in foraging strategies, and second this latter is due to a divergence in whole-genome expression. Using comparative genomics, we showed that the for gene was conserved across insects considering both sequence as well as gene model complexity. Polymorphism analysis did not support the occurrence of two allelic variants diverging across the two populations, yet asexual population exhibited less polymorphism compared to the sexual one. Sexual and asexual transcriptomes sharply split up, with 10.9% of differentially expressed genes, but these were not enriched in behavioural related genes. We showed that the for gene was more expressed in asexual female heads than in sexual ones, and that asexuals were the ones that explored more the environment and exploited more host patches. Overall, these results suggested that a fine tuning in the for gene expression between populations may have led to distinct foraging behaviours. We hypothesized that reproductive polymorphism and coexistence in sympatry of sexual and asexual populations specialized to different ecological niches via divergent optima on phenotypic traits, could imply adaptation through different expression patterns of the for gene and at many other loci throughout the genome.

The *for* gene as one of the drivers of foraging variations in a parasitic wasp

Genomic bases underlying foraging in parasitoid

Aurore Gallot[1]*, Emmanuel Desouhant[1], Vincent Lhuillier[1], David Lepetit[1], Adil El Filali[1], Laurence Mouton[1], Cristina Vieira-Heddi[1], Isabelle Amat[1].

[1] Université de Lyon, Université Lyon1, LBBE – Laboratoire de Biométrie et Biologie Evolutive – UMR 5558, CNRS, Villeurbanne, France

*Corresponding author:*aurore.gallot@univ-lyon1.fr*, https://orcid.org/ **0000-0003-4552-090X**

ABSTRACT

Foraging behaviours encompass strategies to locate resources and to exploit them. In many taxa these behaviours are controlled by a major gene called *for,* but mechanisms vary between species. In the parasitoid

wasp *Venturia canescens* , sexual and asexual populations coexist in sympatry but differ in their foraging behaviours. Here we explored the molecular bases underpinning this divergence in foraging behaviours by testing two mutually non-exclusive hypotheses: firstly the divergence in the *for* gene results in difference in foraging strategies, and second this latter is due to a divergence in whole-genome expression. Using comparative genomics, we showed that the*for* gene was conserved across insects considering both sequence as well as gene model complexity. Polymorphism analysis did not support the occurrence of two allelic variants diverging across the two populations, yet asexual population exhibited less polymorphism compared to the sexual one. Sexual and asexual transcriptomes sharply split up, with 10.9% of differentially expressed genes, but these were not enriched in behavioural related genes. We showed that the *for*gene was more expressed in asexual female heads than in sexual ones, and that asexuals were the ones that explored more the environment and exploited more host patches. Overall, these results suggested that a fine tuning in the *for* gene expression between populations may have led to distinct foraging behaviours. We hypothesized that reproductive polymorphism and coexistence in sympatry of sexual and asexual populations specialized to different ecological niches*via* divergent optima on phenotypic traits, could imply adaptation through different expression patterns of the *for* gene and at many other loci throughout the genome.

Keywords: behavior, foraging, hymenoptera, transcriptomic, populations.

1 INTRODUCTION

Within a population, individuals tend to exhibit similar traits given than sexual reproduction, through meiosis and fecundation, homogenizes genotypes and thus prevent from phenotype divergence. Loss of sexuality have been frequently recorded in a diverse array of eucaryotic taxa with three possible origins leading to the emergence of asexual lineages: mutation, hybridization (Normark, 2003) or endosymbiotic infection (Stouthamer et al., 1990). When reproductive modes are exclusive; asexuals start to diverge by accumulating genetic mutations since they no longer exchange gene flow with the sexual population from which they originated. Competitive interactions should favor one reproductive mode over the other (Lively, 2010). Asexuals have a demographic advantage by producing only females, they avoid the cost of males (Maynard-Smith, 1978). In contrast, sexual populations maintain greater genetic diversity which may confer a decisive advantage in changing environment (Otto, 2009). If geographical or ecological heterogeneity allows ecological specialisation, with each lineage performing better in a specific habitat, sexual and asexual populations can coexist in different geographical areas or sympatrically (Bell, 1982; Lynch, 1984). Such coexistence of sexual and asexual lineages is called reproductive polymorphism. In this case, adaptations in behavior, morphology or life history traits, should distinguish sexual from asexual populations.

Such cases of reproductive polymorphisms have been widely reported in haplodiploid arthropods. Van der Kooi et al. described in an exhaustive database 765 parthenogenetic species, among which 143 presented evidence of sexual lineages as well, representing 19% of the described species having reproductive polymorphism (van der Kooi et al., 2017). Although this proportion is likely to be underestimated, this nevertheless indicates the frequent occurrence of the coexistence between sexual and asexual populations. Thus, the competitive exclusion of one or the other reproductive modes is not the norm, the ecological conditions allowing their coexistence seems to be frequently met. In the hymenopteran parasitoid *Venturia canescens* (Ichneumonoidea : Gravenhorst) asexual populations coexist in sympatry with sexual populations (Beukeboom et al., 1999; Schneider et al., 2002) each better adapted to specific ecological niches (Amat et al., 2017). In particular, divergence in the foraging behaviours have been shown in many population pairs coming from different localities by a meta-analysis (Amat et al., 2017). Asexual wasps have a better capacity than sexual wasps to exploit their environment given their higher capacity to find hosts (Liu et al., 2009), a larger egg load and their ability to lay eggs faster (Pelosse et al., 2007). In contrast, sexual wasps explore better their environment than asexual wasps as they flight longer and faster, have higher energy content, and live longer (Lukáš et al., 2010). *V. canescens* asexual reproduction involves a central fusion automictic parthenogenesis, which means that some genetic recombination occurs during the early stages of oogenesis. Thus, genetic variation still exists between asexual offspring's even if an irreversible increase of homozygosity in populations occurs over time (Beukeboom and Pijnacker, 2000; Mateo Leach et al., 2009). The divergence in

foraging behaviours observed between sexual and asexual wasps *V. canescens* is certainly based on genetic divergence between the two populations, since no more genetic exchanges through mating occurred in natural populations (Mateo Leach et al., 2012).

Genetic control of foraging behaviours has been mainly studied in*Drosophila melanogaster* where two distinct types of strategies have been characterized (Allen et al., 2017; Anreiter et al., 2017; de Belle and Sokolowski, 1989; Osborne et al., 1997; Sokolowski, 1980). While sitters hug the boundaries of a food patch but remain focus within one food patch, rovers travel greater distances within and between food patches, thus exploring widely their environment and exploiting more food resources (Sokolowski, 1980). The two strategies are under the control of a major gene called *foraging (for)* . Despite the many genes involved in generating foraging behaviours (Anreiter et al., 2017), manipulations of the *for* expression are sufficient to modify them (Osborn et al., 1997). The control of foraging behaviours by the *for* gene has been demonstrated in drosophila larvae as well as in adults, and encompass both searching for food resources and oviposition sites (Edelsparre et al., 2014; McConnell and Fitzpatrick, 2017). The *for* gene has two allelic variants conserved during evolution: rovers have at least one dominant allele (*for* ^{R}), while sitters have the two recessive alleles (*for* ^{S}). Both alternative behaviours are maintained by selection; patchy food and high population densities advantage rovers, while evenly distributed food and low population densities advantage sitters (Sokolowski et al., 1997). Genotype differences between rovers and sitters are reflected by differences in the *for* gene expression, higher in rover heads than in sitters, and so is the enzymatic activity of the corresponding protein PKG (Osborn et al., 1997). The role of the *for* gene as a single major gene influencing foraging behaviours has been characterized in many animal species. Orthologs of the *D. melanogaster for* gene influence foraging behaviours in taxa as diverse as nematodes (Hao et al., 2011; Hong et al., 2008) or mammals (Struk et al., 2019). Most of the studies on genetic control of foraging behaviours has been conducted on insects species, such as diptera (*D. melanogaster* , *Aedes aegypti* (Keating et al., 2013)), hymenoptera *(Apis mellifera*(Ben-Shahar et al., 2002), *Bombus terrestris* (Tobback et al., 2011), *Pheidole pallidula* (Lucas and Sokolowski, 2009),*Pogonomyrmex barbatus* (Ingram et al., 2005), *Vespula vulgaris* (Wenseleers et al., 2008), orthoptera (*Schistocerca gregaria* (Lucas et al., 2010)), lepidoptera (*Sesamia nonagrioides* (Chardonnet et al., 2014)). In all those species, the*for* gene contributes to foraging behaviours. However, the existence of allelic variants as well as the relationships between the*for* expression level and foraging behaviours exhibit variations between species. For example, within Hymenoptera, the eusocial honey bee (*A. mellifera*) displays caste division where young workers take care of the hive, whereas older workers of the colony are foragers. Albeit only one allele of *for* has been identified in this species, foragers exhibit higher gene expression associated with a higher corresponding PKG activity (Ben-Shahar, 2005). A same overexpression of the *for* gene have been shown in *B. terrestris* (Tobback et al., 2011), phylogenetically related to honey bee. In contrast, *for* expression was found to be higher in nurses than in foragers in three other eusocial hymenopteran species with a same age-dependent division of labor: the ants (*P. barbatus* ,*P. pallidula*), and the common wasp (*V. vulgaris*) (Ingram et al., 2005; Lucas and Sokolowski, 2009; Wenseleers et al., 2008). Although influence of foraging behaviors by the *for* gene has been maintained during evolution, the underpinning mechanisms variate with opposite patterns within eusocial Hymenoptera species. Studying the genetics of the foraging behaviours on additional hymenopteran species would make it possible a better understanding of their evolution in this taxon. *V. canescens* that belongs to the Ichneumonoidea, a parasitoid superfamily basal within the Apocrita group which include all other eusocial hymenopteran species (bees, ants and wasps) (Peters et al., 2017), appears to be a relevant model for studying the genetic bases underlying the variability of foraging behaviours and their evolution within Hymenoptera.

Here, we investigated the genetic bases underpinnings the variability of foraging behaviours observed between sexual and asexual *V. canescens* populations. We explored two non-mutually exclusive hypotheses that could explained the divergence observed in foraging behaviours: firstly, a divergence in the *for* gene, and second a divergence in whole-genome expression. We first proceed to the characterization of the *for* gene in the *V. canescens* using genomic and transcriptomic sequences: 1) we described the *for*orthologous and reconstructed its evolution in insects; 2) we annotated the full gene model by analyzing sexual and asexual transcriptomes; 3) we described allelic variations in sexual and asexual populations. We explored the second hypothesis

by studying the differential gene expression between sexual and asexual populations, with a special focus on the behavioral genes. Finally, we tested whether variations in the*for* gene expression could explain the observed variations in foraging behaviours between sexual and asexual populations, by coupling a behavioral experiment with the *for* gene quantification.

2 MATERIALS AND METHODS

2.1 Field sampling and insect rearing

V. canescens is a solitary endoparasitoid of caterpillars of pyralid moths (Salt, 1976). The females used in the experiments come from sexual and asexual populations collected annually near Valence (N44°58'21", E4°55'39"). In this unique location, individuals of the sexual population were usually sampled in an orchard, while individuals of asexual population were mostly collected close to grain silos. Caterpillars of *Ephestia kuehniella* (Zeller) were left a week exposed to parasitoids, then brought back to the laboratory waiting the emergence of parasitoids. Virgin emerging *V. canescens* females were isolated and left with hosts in order to sexed their progeny. In*V. canescens* sex determination is haplodiploid: sexual females have a parthenogenetic arrhenotokous reproduction, i.e. unfertilized eggs produce haploid males while diploid females resulted from fertilized eggs. Thus, virgin arrhenotokous females produce only males. In contrast, virgin asexual thelytokous females produce only females. Sexual and asexual wasps were maintained separately on the host *E. kuehniella* feed on semolina, were they produced kairomones attracting for parasitoids from mandibular gland secretions (Castelo et al., 2003). Insects were grown under constant environment (25 ± 1°C, 55 ± 5 % RH, 12:12 LD).

2.2 Annotation of the *for* gene in *V. canescens* genome

Orthologs, *i.e.* genes descended from the same ancestral sequence separated by a speciation event, often have the same function, hence we first searched the orthologs of the *for* gene in the genome of*V. canescens* . A set of 40 orthologs *for* sequences from 38 insect species, as well as the branchiopoda *Daphnia pulex* and the mouse (*Mus musculus*) sequences, were identified using ortholog annotation in EnsemblMetazoa database and literature (table S1). To identify *for* ortholog in *V. canescens* genome, we used the reciprocal best hits with tblastn with a set of 42 *for* orthologs protein sequences previously described as query, and the *V. canescens* genome as database (*http://bipaa.genouest.org/sp/venturia_ canescens/*V.1.0). The *for* gene that was localized on the scaffold 64, contained the longest open reading frame (*Vcan27709*) constituted of 2,445 nucleotides encoding 815 amino acids.

2.3 The *for* gene phylogenetic reconstruction in insects

The putative *V. canescens for* sequence was added to the set of 42 orthologs *for* , then aligned using MUSCLE (Edgar, 2004). The corresponding protein alignment diverged in N-terminal but was conserved in C-terminal. Alignment was manually curated, most conserved residues were selected using Gblock, the resulting alignment consisted of 533 amino acids. To reconstruct the *for* phylogeny and position the *V. canescens for* among other insect sequences, ProtTest v3.4.2. was used to determine the best-fit model of protein evolution using AIC (Abascal et al., 2005). JTT model of protein evolution was used and topology optimization was carried out using best of NNI and SPR options. The phylogenetic tree was constructed with maximum likelihood method using PhyML implemented in Seaview (v 4.7) (Gouy et al., 2010). Default aLRT (SH-like) was used for branch support (Anisimova and Gascuel, 2006).

2.4 RNA extraction and sequencing

A total of 6 RNA-seq libraries were prepared: sexual and asexual populations were constituted both by 3 biological replicates. Each replicate was constituted of a pool of 30 individual heads taken from emerging females and flash frozen, next used as an input for RNA extraction. Heads were crushed using steel beads and Qiagen TissueLyser (45s, 25hz). Total RNAs were extracted using Rneasy Mini Kit (Qiagen) following the manufacturer protocol and including the DNase step. RNA integrity was controlled using gel electrophoresis and quantified with Nanodrop. After integrity control and quantification, polyadenylated RNAs were enriched from 1 µg of high-quality total RNAs with oligo-dT magnetic beads, then fragmented and converted to cDNAs (Illumina TruSeq Stranded mRNA Library Prep kit). Fragments around 200bp were selected, adapters

ligated, and fragments amplified by PCR to generate DNA colonies. Each library was labelled, multiplexed and pooled for sequencing on a HiSeq 2500 Illumina sequencer (Fasteris, Switzerland), with a paired-end protocol (2x150bp).

2.5 The *for* gene model reconstruction

We identified all the isoforms of *for* transcript and reconstructed the *for* gene model in *V. canescens* by screening the 6 RNA-seq libraries from sexual and asexual populations and focusing on the *for* reads. KisSplice 2.5.4. (Sacomoto et al., 2012) is a method based on De Bruijn graphs that allows identification of all variants without using a reference genome, including single nucleotide polymorphism (SNPs), indels and alternative splicing events. In parallel, we built a *de novo* transcriptome assembly with Trinity (Haas et al., 2013).

2.6 Polymorphism analysis at the *for* locus

To evaluate the *for* gene polymorphism, we localized all the SNPs based on the list of all the SNPs, insertions and deletions across all the *for* isoforms, previously identified with KisSplice. We then used KisSplice2RefTranscriptome to position each SNP on *for*isoforms. Finally, we used the R package KissDE in order to find SNPs that significantly differed in frequency across sexual and asexual populations (adjusted P -values<0.05).

2.7 Differential expression analysis

The genome-wide divergence between sexual and asexual populations was estimated using RNA-seq libraries to identify differentially expressed genes (DEGs) between the two populations. Reads quality was first assessed with FastQC, then reads were trimmed and filtered using trimmomatic with minimum length set to 75pb. After filtering, the transcriptomic dataset included a total of 91 millions of reads of which an average of 92% were successfully aligned on *V. canescens*transcriptome using HiSat2 (Kim et al., 2019) (table S2). Genes with differential expression between sexual and asexual populations were identified using negative binomial GLM implemented in the program DESeq2 (Love et al., 2014). We tested for differential expression of all transcripts with an average level of expression superior to 10 reads per gene (n =14,106). A gene was considered differentially expressed (DE) when the false discovery rate (FDR) adjusted p- value was inferior to 0.05, without applying supplementary fold change threshold.

2.8 Functional analysis

The *de novo* transcriptome was annotated using BLAST and Gene Ontology tools to assign biological function to transcripts. Then, we focused our analysis on the transcripts related to the 'behaviour' GO term or to any of its child terms, thus annotating a functional group of transcripts related to behaviour. This list of transcripts was crossed with the previously established lists of DEGs between the two populations.

2.9 Behavioural experiment

We set up an experimental design to quantify exploitation and exploration of host patches by individuals from sexual and asexual populations. Experimental device contained two host patches placed 20 cm far from each other inside a box (50 x 16 x 8 cm) with 2 side holes covered with veils to allow ventilation. The two host patches were made of Petri dish (5.5 cm Ø) containing six 21-days old larvae of *E. kuehniella* and semolina to the rim, prepared seven days before the test and covered with a thin gauze to prevent larvae from escaping. Each host patch was embedded in clean semolina in the middle of a bigger Petri dish (13 cm Ø). Every morning, wasps were collected at the emergence and placed individually in tubes with one drop of water. The day after, males and females of the sexual strain were gathered in a cage to mate. Females were observed and gradually picked up in a tube as they have mated until the behavioural experiment. In the meantime, emerging asexual females were placed in another cage in the same conditions. At the beginning of the experiment a single female was inserted in the box and deposited in the middle of the left patch, called 'patch 1', whereas the right patch was called 'patch 2'. A total of 34 females (17 asexuals, 17 sexuals) were tested in random order. Foraging behaviours were followed during 20 min by recording 4 metrics

with Jwatcher (Blumstein and Daniel, 2007): i) probing, wasps probed the substrate with ovipositor once presence of hosts detected thanks to kairomones; ii) cocking, a peculiar movement of the abdomen observed after egg laying, when the female load of a new egg at the tip of its ovipositor (Rogers, 1972); iii) moving outside of patches (*i.e.* flying or walking); and iv) time dedicated to hosts, next called patch residence time (PRT) *i.e.* sum of the time spent on patch 1 and patch 2. The female was considered to have left a patch when more than 150 seconds was spent outside of the patch, hence PRT includes short excursions outside patch boundaries. Exploitation was considered as the capacity of females to find hosts within patches. Host patches exploitation has been measured by the means of 2 parameters: i) total PRT was used as a synthetic parameter to summarize the exploitation of the two patches (sum of PRT on patches 1 and 2); ii) the total number of ovipositions,*i.e.* the number of cockings, used to assess the success of exploitation. Exploration was considered as the ability to visit the entire experimental device (i.e. environment) and has been quantified using two parameters: i) the proportion of females that manage visiting the two host patches, considered as the aptitude to locate new resources; ii) the number of switches between the two host patches, considered as the ability to navigate between different resources. Immediately after behavioural experiment, all the 34 wasp heads were individually collected to quantify expression of the *for* gene, while abdomens were dissected to count the number of eggs in the ovarioles, next called egg load. Heads were stored on ice in 10 µl of RNA-later (Sigma-Aldrich), then at -20°C until RNA extraction.

2.10 Quantification of the *for* gene expression

We quantified the *for* gene in each individual wasp head using RT-qPCR to be able to correlate foraging behaviours with the *for*gene expression. Once all samples collected, the 34 RNA extractions were performed in one batch by series of 12 randomized samples, using the protocol described above. First-strand cDNA was synthesized from 70ng of total RNA using SuperScript III first strand synthesis system (ThermoFisher scientific) with random hexamer primers, and followed by a RNAse-H step. Quantification was conducted on the *for* gene, together with two reference genes (*rpl32* and *gapdh*) used for normalization between samples to control variations in extraction yield, reverse transcription yield, efficiency of amplification. Reactions were performed on a CFX-96 (BioRad) using 1:10 diluted cDNA and SYBR Green master mix (BioRad), according to the manufacturer instructions. Amplification conditions were a first step of denaturation (95°C, 1 min) followed by 40 cycles of denaturation (95°C, 10 sec) and elongation (melting temperature, 30 sec). Details on primers and melting temperatures were listed in table S3. Fluorescence was quantified at the end of each cycle, and the quantification cycle (C_q) corresponding to the start of exponential phase amplification was measured. Each sample was quantified twice: all duplicated C_q values varied less than 0.5 cycle, indicating an elevated replicability. The expression level of the *for* gene was determined relatively to the expression level of both reference genes*rpl32* and *gapdh,* using [?][?]C_q method (Livak and Schmittgen, 2001). Results were consistent whatever the reference gene used, and both *rpl32* and *gapdh* provided satisfactory quality control (low C_q values and low variations across samples) (figure S1 and table S4). Therefore, we finally used the mean C_q between *rpl32* and *gapdh* for normalization to increase precision of the results. The *for*expression values were expressed using relative values comparing each individual to the median individual, considered as the value 0. Negative values indicated thus individuals with *for* expression lower than the median, and positive values indicated individuals with *for*expression superior to the median.

2.11 Statistical analysis

The foraging behaviours, decomposed into exploitation and exploration, each measured by a set of parameters previously defined, were analyzed using Generalized Linear Models (GLM). The population (sexual or asexual), the *for* gene expression (fold change), the egg load (a proxy for parasitoid fitness (West et al., 1996)), and the double interactions with the variable population were included as predictor variables. PRT was analyzed with GLM with a Gamma distribution for errors and inverse link. Number of switches between host patches was analyzed with GLM with a Poisson distribution and log link. Number of cockings was also analyzed with GLM with a Poisson distribution for error and log link. PRT was added to the full model since cocking probability increase with PRT. The *for* gene expression was analyzed with linear model using population, egg load and their interaction as explanatory variables. Least contributive variables in all

models were iteratively removed using backward selection to select optimal models. All statistical analyses were performed with R (R Core Team, 2017).

3 RESULTS

3.1 Identification of the *for* gene in *V. canescens* genome and *for* gene evolution in insects

We identified a sequence candidate to be *for* ortholog in the *V. canescens* genome, and then aligned it to a set of *for* orthologs in order to reconstruct the evolutionary history of the *for* gene. The resulting maximum likelihood tree robustly related the major represented insect clades: hymenoptera, orthoptera, coleoptera, diptera, lepidoptera (figure 1A). A majority of one-to-one orthologous relations was detected, with the exception of rare duplication events. All ten sequences from hymenopteran species constituted a monophyletic group highly supported (bootstrap value $> 95\%$, figure 1A). Within this group, *V. canescens* (Ichneumonidae) clustered with *Nasonia vitripennis* (Chalcidoidea) to constitute the parasitoida group. The phylogenetic reconstruction confirmed that one unique sequence within the *V. canescens* genome was ortholog to the *for* gene in *D. melanogaster* , and was then annotated as the *V. canescens for* gene (*Vcan_for)* .

3.2 Characterization of the *for* gene model in *V. canescens*

We produced RNA-seq libraries from sexual and asexual populations with a triple objective: i) reconstruct the *for* gene model, *i.e.* the region of the gene that is supposed to be transcribed into RNA; ii) evaluate polymorphism at the *for* locus within sexual and asexual populations, and iii) assess genome-wide differences in gene expression between the two populations. To produce an accurate model of the *for* gene, we thus screened the RNA-seq libraries from sexual and asexual female heads searching for all reads mapping on this locus and reconstructed all the transcripts. We identified four separate transcription start sites, supporting a gene model that contains four independent promoters (*pr1 -pr4*) corresponding to four distinct open reading frames (figure 1B). The longest open reading frame *Vcan27709* started with *pr1* , exhibited 7 isoforms that mainly differed in their untranslated regions (UTRs). *Vcan27708* transcript started with *pr2* and showed 4 isoforms, while *Vcan27707* (*pr3*) possessed 5 isoforms. Finally, the shortest transcript *Vcan27706* starting with *pr4* presented one isoform. Overall, a total of 13 exons were identified whose different combinations constituted 17 different isoforms. The 9 first exons exhibited alternative splicing, thus isoforms essentially differed in their 5' UTR and the corresponding N-terminal coding sequences. In contrast, the last 4 exons were constitutive of all isoforms (excepted *Vcan27707_i12, Vcan27709_i14, Vcan27708_i17;* that contained an early stop codon), and constituted one unique 3' extremity (figure 1B) encoding the C-terminal part of the protein, containing the two cGMP-binding domains as well as the kinase domain.

3.3 Allelic variation at the *for* locus between sexual and asexual populations

We focused on the population polymorphism at the *for* locus, by screening RNA-seq libraries based on 90 females from sexual and asexual populations, we identified a total of 15 single nucleotide polymorphim (SNP) (table 1). Only 3 SNPs were located within the coding region, including 2 synonymous SNPs and one single non-synonymous mutation. The 12 remaining SNPs were located outside of coding sequences, within UTRs. Among the 15 SNPs, 14 variants exhibited significant differences in frequency between sexual and asexual populations (table 1). All those variants were polymorphic in sexual population, while nine were fixed in asexual population. Together, these results do not support the existence of two allelic variants differing between sexual and asexual populations. We rather described a variety of polymorphic sites accumulated all along the locus, with an important reduction of polymorphism detected in the asexual population. Moreover, the protein sequence was little affected by polymorphism, with only one non synonymous variant recorded and located outside of the functional sites. Nonetheless, the numerous polymorphic sites reported all along the *for* gene could affect the transcription or the alternative splicing of the gene rather than the sequence of the encoded protein itself.

3.4 Genome-wide expression divergence across sexual and asexual populations

Overall, we found that gene expression strongly diverged according to sexual or asexual population. The principal component analysis based on the expression of all the genes showed that the first axis separated

sexual from asexual population and explained 82% of the total variance (figure 2A). Among the 14,106 transcripts that passed the expression filter, a total of 1,539 genes were DE (P -adj<0.05) between sexual and asexual population, representing 10.9% of the transcriptome. The *for* transcript, represented in the transcriptome by its longest isoform (*Vcan27709*transcript*)* , was not included within this list of DEG (rank 2,507/14,106, P -adj = 0.168) (figure 2B). Although the$P$ -adj value being above the significance level, the analysis of normalized counts of the *for* transcript across the 6 libraries showed that *for* expression was about 10% higher in asexual compared to sexual population (figure 2C). None of the 3 others*for* transcripts (*Vcan27706* , *Vcan27708* and*Vcan27709)* exhibited significant differential expression between sexual and asexual populations, but all showed the same expression pattern (figure S2).

3.5 Behavioural genes expression divergence between sexual and asexual populations

The *de novo* transcriptome assembly was constituted of a total of 22,333 transcripts that were annotated using BLAST and Gene Ontology tools. Among those, 18,316 get a blast hit (82%) and 12,923 get at least one GO term annotation (58%). We selected the 'behaviour' Gene Ontology term, as well as all its child related GO terms. In this way, we annotated 249 transcripts with putative functions associated with behavior in *V. canescens* . Among them, we reported 26 transcripts that were DE between the two populations, which represented potential candidates in the differences in foraging behaviours observed between sexuals and asexuals (table 2). The proportion of behavioural genes with differential expression between the 2 populations was not different compared to the full transcriptome one (26/249 *vs* 1,539/12,567; χ^2= 0.44, P -val=0.50). Among those, we noticed a majority of transcripts related to sensory behaviour: chemosensory (18 transcripts) or visual (2 transcripts). The other functions detected were the locomotory behaviour (2 transcripts), learning and memory (2 transcripts), reproductive behaviour (1 transcripts) and rhythmic behaviour (1 transcript).

3.6 Asexual females exploited more hosts and explored more environment than sexual females

In the behavioural experiment asexual wasps exploited more hosts than sexual ones, by allocating more time to hosts (figure 3A; χ^2= 3.81, df=1, P <0.01). On average, asexual females spent twice more time on host patches compared to sexuals (655.5 ± 60.8s *vs* 333.3 ± 80.4s). Neither the egg load, the *for* expression nor their interactions with population were significantly different. Asexual females laid twice more eggs than sexual ones (figure 3B; χ^2=4.94, df=1,P <0.05), with on average 1.24 eggs laid by asexuals (± 0.32) compared to 0.53 eggs by sexuals (±0.17). Time spent on a patch determined the number of eggs laid since PRT has positive effect on the number of cockings (χ^2=14.09, df=1,P <0.001). However, *for* expression has a marginal, though not statistically significant, effect on the number of cockings (χ^2= 3.01, df=1, P =0.08), with a number of cockings increasing in individuals with higher *for*expression. The interactions between *for* expression and population, and between *for* expression and PRT, did not explain the number of cockings. Asexual females also explored more the environment than sexual ones. We did not detect differences between sexual and asexual populations in the proportion of females finding the second hosts patch (11/17 *vs* 14/17 respectively, Fisher exact test, P =0.44). However, asexual females switched more from one patch to another than sexual females (figure 3C; χ^2=4.937, df=1, P <0.05), with on average 3-fold more changes in asexual females (2.12 ± 0.37 in asexuals*vs* 0.71 ± 0.14 in sexual females). Switches between host patches were not influenced by other variables, nor by their interactions with population. Together these results showed that asexual females exploited more hosts, with more time spent on host patches and more eggs laid, and explored more the environment by changing more often of host patches.

3.7 Expression of the *for* gene and correlations with behaviours in sexual and asexual females

The *for* gene expression was superior in asexual female heads (1.06±0.60) compared to sexual ones (-0.98±0.54) (figure 4A; F=7.62, df=1 and 31, P <0.01). Within each population,*for* expression decreased with egg load (figure 4B; F=6.7, df=1 and 31, P <0.05). There was no significant interaction between egg load and population on the *for* gene expression. When analyzing both populations separately, the number of cockings increased with PRT in asexual females (χ^2=6.00, df=1,P =0.014) but decreased with *for* expression (χ^2=4.96, df=1, P =0.026) (figure 4C). In sexual females, number of cockings is correlated with PRT

(χ^2=8.1061, df=1, P =0.004) but not with*for* expression.

4 DISCUSSION

The *for* gene exhibited a strong sequence conservation across insects, consistent with the function conservation in influencing the foraging behaviours described in numerous insects (Reaume and Sokolowski, 2009). Beyond *for* sequence conservation, we also showed a conservation of the *for* gene model complexity between *V. canescens* and *D. melanogaster* with four alternative promoters encoding four proteins differing in their N-termini (Allen et al., 2017). The use of alternative promoters represents a source of diversity and flexibility in the regulation of gene expression, and ultimately function. This as has been particularly demonstrated in the*for* gene, whose promoter variations cause changes in both tissue localization and substrate specificity. Indeed, *pr1-for* and*pr4-for* transcripts were expressed within neurons, while*pr2-for* and *pr3-for* transcripts were localized in glia cells of fruit flies central nervous system (Allen et al., 2018; Dason et al., 2020). The isoform *pr1-for* was presumed to be the only transcript necessary to forage since in mutants *pr1-for*expression in neurons was the only required to rescue larval foraging behaviours (Allen et al., 2018). Variations in the N-termini are critical to the specificity of PKG-substrate interactions (Pearce et al., 2010). PKG phosphorylate serine and threonine residues on a dozen of proteins known to modulate muscle activity and neuronal signaling pathways (Edelman et al., 1987; Schlossmann and Desch, 2009). Such variety of substrates may explain the pleiotropic effects of the*for* gene. Conservation of gene model complexity between the diptera *D. melanogaster* and the hymenoptera *V. canescens*supports the importance of maintaining such complexity to regulate alternative foraging behaviours. However, our work did not allow the characterization of qualitative differences in *for* isoforms between sexual and asexual populations, but rather suggested a decrease in all *for* isoforms transcription in the sexual population.

Polymorphism analysis revealed 15 SNPs along the *for* gene, most of them varying in frequency across populations, and supported a major reduction of genetic diversity that occurred in asexuals rather than the presence of two allelic variants diverging between sexual and asexual populations. Such reduction of polymorphism in asexuals was expected: in general thelytokous individuals are more homozygous than arrhenotokous ones (Beukeboom and Pijnacker, 2000) and this has been already shown in *V. canescens* with a study based on 15 microsatellites, that were all homozygous (Mateo Leach et al., 2012). In contrast, some genetic diversity still persisted at the *for* locus in asexuals. The vast majority of identified SNPs did not affect the protein sequence itself since occurring outside of coding region or corresponding to synonymous polymorphism. A single SNP corresponding to a non-synonymous mutation was located at the N-terminal part of the predicted PKG I, corresponding to the substrate binding region of the protein, outsides the kinase and cGMP binding domains. By comparison, *rover* and *sitter*alleles differed by more than 300 SNPs segregating in *D. melanogaster* , but also involved regulatory mutations rather than changes in aminoacid sequence (Allen et al., 2017). In contrast, the two allelic variants identified in the moth *Sesamia nonagrioides*differed by only one non-synonymous SNP located within the kinase domain of the protein, each variant was associated with different levels of*for* expression, PKG activity, and distinct behaviours (Chardonnet et al., 2014). However, previous studies in hymenopteran species have not shown any evidence of the existence of allelic variants at the*for* locus, neither did the present study in *V. canescens* .

Differences in foraging behaviours recorded between sexuals and asexuals should rely on genome divergence since there is no more gene flows between the 2 populations (Mateo Leach et al., 2012). Previous study revealed such genome divergence since individuals from sexual and asexual populations can be distinguished based on microsatellites. However, how much gene expression diverged at the genome-wide scale between the two populations has not been studied so far. By comparing head transcriptomes, we reported that the 2 populations clearly split up, with a total of 1,539 DEGs. This proportion of 11% of DEGs between 2 populations from a same species is high and almost as high as that observed in recently diverged species such as *Drosphila pseudoobscura pseudoobscura* and *D. pseudoobscura bogotana*(˜0.25mya divergence and 14.6% DEGs) (Gomes and Civetta, 2015). While behavioural divergences are among the most remarkable differences between sexual and asexual wasps, the behaviour-annotated group of genes was not overrepresented within DEGs. Among the behavioural genes whose expression varies between populations, transcripts involved

in sensory perception (olfactory, sensitive, visual) were the most numerous. Chemosensory genes evolved rapidly and played important role in adaptation (Brand et al., 2015). While the*for* gene has been implicated in foraging behaviours in a variety of organisms, none of the *for* isoforms were detected as DE. Analysis of RNA-seq data at this locus showed that all isoforms of the*for* gene were more expressed in asexual population than in sexual population, although beyond significance threshold.

Globally the behavioural results were congruent with our predictions, that is asexuals exploited more host patches since they are faster to choose and walk between hosts, and have a greater egg load (Amat et al., 2017). On the opposite, sexual females are better dispersers with both longer and faster flights, they have higher longevity, and greater energy content (Amat et al., 2017). They were thus expected to explore more their environment, yet this prediction was not verified in the current results. The discrepancy could came from the experimental device that might be too small for all exploration-related behaviours to be expressed, in particular dispersal involving long flights with high energy costs (Amat et al., 2012). Field experiments conducted with*D. melanogaster* showed that rovers exhibit higher dispersion with both greater dispersal tendencies and longer distances of flight than sitter flies, and that artificial increasing of *for* expression in the brain and nervous system increases dispersal in sitters (Edelsparre et al., 2014). Our study, conducted in the lab, does not allow assessing wasps dispersal ability. Nevertheless, by showing that asexual females switched more frequently between host patches than sexual ones, the experimental device is relevant to detect differences in some aspects of the exploration between the two populations.

V. canescens asexual females present homology with the rover phenotype observed in *D. melanogaster* both by exploring and exploiting more. The *for* gene is more expressed in asexual wasps than in sexual ones, consistently with the *Drosophila* rover model. Previous classification of the *for* transcript among non-DEG might be due to the low number of RNA-seq replicates (3), while RT-qPCR was conducted on a higher number of individuals (17) thus increasing statistical power in the detection of DEG. The higher fecundity is another common characteristic between *D. melanogaster* rovers (McConnell and Fitzpatrick, 2017) and *V. canescens* asexuals, here measured by both a higher egg load, corresponding to their potential fitness, and higher number of eggs laid, corresponding to their effective fitness but measured during a short period. Hence in these two species, individuals that exploited and explored more were also the more fecund, and the ones with the higher*for* expression. A major contribution of the present study is the joint analysis of the *for* expression and foraging behaviours measured at the individual scale, that conferred information on the inter-individual variations and allowed studying correlations between these traits, beyond average measures. Two major results emerged from this approach: the first one is that the egg load decreases in females with the highest *for* expression; and the second is a decrease of eggs laid by asexual females with the highest *for* expression. These two correlations were consistent and suggested that an increase in the *for* expression may be costly for females and could result in a decrease of progeny number. In the wasp *V. canescens* , the cost of reproduction is mostly based on finding hosts to lay eggs, as the egg itself contains little reserve and is not costly to produce (Pelosse et al., 2011). In this case, rather than an energetic cost due to the*for* expression that would directly induce a decrease in fecundity, the cost might be indirect and related to the numerous other functions fulfilled by the highly pleiotropic *for* gene apart from resource searching behaviours, such as learning, memory, or social interactions (Alwash et al., 2021; Reaume and Sokolowski, 2009).

Given the extent of transcriptomic divergence, with hundreds of DEGs between sexual and asexual populations, and in the absence of functional analysis, we cannot firmly conclude on the functional role played by the slight differences in the *for* expression recorded in the differences of foraging behaviours. It is worth mentioning that comparison between rovers and sitters in *D. melanogaster* showed that differences in the *for* expression were small but consistent (Osborne et al., 1997). Drosophila rovers and sitters have shown differences between their transcriptomes, apart from the single variation in the *for* expression (Kent et al., 2009). Honey bee nurses and foragers differed by about 40% of their brain transcriptome (Whitfield et al., 2003). However, manipulation of the *for* gene expression or the corresponding PKG enzyme activity was sufficient to modify foraging behaviours in the two species (Ben-Shahar et al., 2002; Osborne et al., 1997). Therefore, the differences in the *for* gene expression detected in the current study between sexuals and asexuals, although moderate, might nevertheless have an essential function in the differences in foraging

behaviours reported between *V. canescens*populations.

The present study that highlights the molecular bases underpinning the variability in foraging behaviours in the parasitoid wasp *V. canescens* brings insights on the evolution of the control of foraging behaviours by the *for* gene in hymenoptera. So far, studies have focused on social hymenopteran species that acquired eusociality independently: bees, ants and wasps. All these studies revealed a caste-specific *for* expression correlated with foraging intensity but with opposite patterns: honey bee (*A. mellifera*) and bumblebee (*B. terrestris*) foragers exhibit a higher *for*expression compared to nurses (Ben-Shahar et al., 2002; Tobback et al., 2011), whereas nurses presented a higher *for* expression compared to foragers in ants (*P. barbatus, P. pallidula*) and common wasp*V. vulgaris* (Ingram et al., 2005; Lucas and Sokolowski, 2009; Wenseleers et al., 2008). The *for* gene influences social behavior in a variety of species and has been postulated to be part of a genetic toolkit involved in the evolution of eusocial insects (Rittschof and Robinson, 2016): whereas acquisition of eusociality relies on the emergence of a forager caste specialized on foraging tasks that appears to be related to differences in the *for* gene expression. In contrast, selection in bees shows opposite patterns to the one described in the ancestral groups of ants and wasps. Parasitoids wasps, which are ancestral to the Apocrita group that includes all social hymenopteran (Peters et al., 2017), are solitary species and therefore do not have foragers. This study suggests that differences in the *for* expression pattern underlying changes in foraging strategies, could be ancestral to Apocrita and precede the acquisition of sociality. In this group, variations in the *for* gene expression would not rely on allelic variants. The present work illustrates an original case of a divergence in foraging behaviours that is not based on caste differences but associated with a difference in the *for* expression between populations that also differ in their reproductive mode. However, the adaptations observed in numerous life history traits in the two populations are not limited to the difference in expression of one gene but could involve differences in the optimum expression pattern of several hundred genes.

Acknowledgements

We thank Francois Debias for his help during field capture and Elsa Day for her help in insect rearing and experiment. This work was funded by the Agence Nationale de la Recherche JCJC AVOIDINBRED (ANR-17-CE02-0004-01) attributed to A.G. This work was performed using the computing facilities of the CC LBBE/PRABI.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.

Allen, A., Anreiter, I., Vesterberg, A., Douglas, S., Sokolowski, M., 2018. Pleiotropy of the Drosophila melanogaster foraging gene on larval feeding-related traits. Journal of neurogenetics 32, 256–266.

Allen, A.M., Anreiter, I., Neville, M.C., Sokolowski, M.B., 2017. Feeding-related traits are affected by dosage of the foraging gene in Drosophila melanogaster. Genetics 205, 761–773.

Alwash, N., Allen, A.M., B. Sokolowski, M., Levine, J.D., 2021. The Drosophila melanogaster foraging gene affects social networks. null 35, 249–261. https://doi.org/10.1080/01677063.2021.1936517

Amat, I., Besnard, S., Foray, V., Pelosse, P., Bernstein, C., Desouhant, E., 2012. Fuelling flight in a parasitic wasp: which energetic substrate to use? Ecological Entomology 37, 480–489.

Amat, I., van Alphen, J.J.M., Kacelnik, A., Desouhant, E., Bernstein, C., 2017. Adaptations to different habitats in sexual and asexual populations of parasitoid wasps: a meta-analysis. PeerJ 5, e3699. https://doi.org/10.7717/peerj.3699

Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Systematic biology 55, 539–552.

Anreiter, I., Kramer, J.M., Sokolowski, M.B., 2017. Epigenetic mechanisms modulate differences in Drosophila foraging behavior. Proc Natl Acad Sci USA 114, 12518. https://doi.org/10.1073/pnas.1710770114

Bell, G., 1982. Masterpiece of Nature: The evolution of sexuality.

Ben-Shahar, Y., 2005. The foraging gene, behavioral plasticity, and honeybee division of labor. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 191, 987–994. https://doi.org/10.1007/s00359-005-0025-1

Ben-Shahar, Y., Robichon, A., Sokolowski, M., Robinson, G., 2002. Influence of gene action across different time scales on behavior. Science 296, 741–744.

Beukeboom, L.W., Driessen, G., Luckerhoff, L., Bernstein, C., Lapchin, L., Van Alphen, J.J.M., 1999. Distribution and relatedness of sexual and asexual Venturia canescens (Hymenoptera), in: Proc. Viiith Meeting Exp. Appl. Entomol. Neth. pp. 23–28.

Beukeboom, L.W., Pijnacker, L.P., 2000. Automictic parthenogenesis in the parasitoid Venturia canescens (Hymenoptera: Ichneumonidae) revisited. Genome 43, 939–944. https://doi.org/10.1139/g00-061

Blumstein, D., Daniel, J., 2007. Quantifying behavior the JWatcher way.

Brand, P., Ramírez, S.R., Leese, F., Quezada-Euan, J.J.G., Tollrian, R., Eltz, T., 2015. Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). BMC evolutionary biology 15, 1–16.

Castelo, M.K., Corley, J.C., Desouhant, E., 2003. Conspecific avoidance during foraging in Venturia canescens (Hymenoptera : Ichneumonidae): The roles of host presence and conspecific densities. Journal of Insect Behavior 16, 307–318. https://doi.org/10.1023/A:1023928204793

Chardonnet, F., Capdevielle-Dulac, C., Chouquet, B., Joly, N., Harry, M., Le Ru, B., Silvain, J.-F., Kaiser, L., 2014. Food searching behaviour of a Lepidoptera pest species is modulated by the foraging gene polymorphism. Journal of Experimental Biology 217, 3465–3473.

Dason, J.S., Cheung, A., Anreiter, I., Montemurri, V.A., Allen, A.M., Sokolowski, M.B., 2020. Drosophila melanogaster foraging regulates a nociceptive-like escape behavior through a developmentally plastic sensory circuit. Proc Natl Acad Sci USA 117, 23286. https://doi.org/10.1073/pnas.1820840116

de Belle, J.S., Sokolowski, M.B., 1989. Rover/sitter foraging behavior in Drosophila melanogaster: Genetic localization to chromosome 2L using compound autosomes. Journal of insect behavior 2, 291–299.

Edelman, A.M., Blumenthal, D.K., Krebs, E.G., 1987. Protein serine/threonine kinases. Annual review of biochemistry 56, 567–613.

Edelsparre, A.H., Vesterberg, A., Lim, J.H., Anwari, M., Fitzpatrick, M.J., 2014. Alleles underlying larval foraging behaviour influence adult dispersal in nature. Ecology Letters 17, 333–339.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32, 1792–1797.

Gomes, S., Civetta, A., 2015. Hybrid male sterility and genome-wide misexpression of male reproductive proteases. Scientific Reports 5, 11976. https://doi.org/10.1038/srep11976

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution 27, 221–224.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript

sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hao, Y., Xu, N., Box, A.C., Schaefer, L., Kannan, K., Zhang, Y., Florens, L., Seidel, C., Washburn, M.P., Wiegraebe, W., 2011. Nuclear cGMP-dependent kinase regulates gene expression via activity-dependent recruitment of a conserved histone deacetylase complex. PLoS Genet 7, e1002065.

Hong, R.L., Witte, H., Sommer, R.J., 2008. Natural variation in Pristionchus pacificus insect pheromone attraction involves the protein kinase EGLA-4. Proc. Natl. Acad. Sci. U. S. A. 105, 7779–7784. https://doi.org/10.1073/pnas.0708406105

Ingram, K.K., Oefner, P., Gordon, D.M., 2005. Task-specific expression of the foraging gene in harvester ants. Molecular Ecology 14, 813–818.

Keating, J.A., Bhattacharya, D., Rund, S.S.C., Hoover, S., Dasgupta, R., Lee, S.J., Duffield, G.E., Striker, R., 2013. Mosquito protein kinase g phosphorylates flavivirus ns5 and alters flight behavior in Aedes Aegypti and anopheles Gambiae. Vector-Borne and Zoonotic Diseases 13, 590–600. https://doi.org/10.1089/vbz.2012.1110

Kent, C.F., Daskalchuk, T., Cook, L., Sokolowski, M.B., Greenspan, R.J., 2009. The Drosophila foraging gene mediates adult plasticity and gene-environment interactions in behaviour, metabolites, and gene expression in response to food deprivation. PLoS Genet 5, e1000609–e1000609. https://doi.org/10.1371/journal.pgen.1000609

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

Liu, Y.Q., Thiel, A., Hoffmeister, T.S., 2009. Odor-mediated patch choice in the parasitoid venturia canescens: Temporal decision dynamics. Entomologia Experimentalis et Applicata 132, 110–117. https://doi.org/10.1111/j.1570-7458.2009.00878.x

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$CT method. methods 25, 402–408.

Lively, C.M., 2010. A review of Red Queen models for the persistence of obligate sexual reproduction. Journal of Heredity 101, S13–S20.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lucas, C., Kornfein, R., Chakaborty-Chatterjee, M., Schonfeld, J., Geva, N., Sokolowski, M., Ayali, A., 2010. The locust foraging gene. Archives of insect biochemistry and physiology 74, 52–66.

Lucas, C., Sokolowski, M.B., 2009. Molecular basis for changes in behavioral state in ant social behaviors. PNAS 106, 6351–6356. https://doi.org/10.1073/pnas.0809463106

Lukáš, J., Bernstein, C., Gu, H., Dorn, S., 2010. Could different environmental conditions select for different flight characteristics in sexual and asexual parasitoid Venturia canescens? Entomologia Experimentalis et Applicata 136, 80–88. https://doi.org/10.1111/j.1570-7458.2010.01000.x

Lynch, M., 1984. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. The Quarterly Review of Biology 59, 257–290.

Mateo Leach, I., Ferber, S., van de Zande, L., Beukeboom, L.W., 2012. Genetic variability of arrhenotokous and thelytokous Venturia canescens (Hymenoptera). Genetica 140, 53–63. https://doi.org/10.1007/s10709-012-9657-6

Mateo Leach, I., Pannebakker, B.A., Schneider, M.V., Driessen, G., Zande, L. van de, Beukeboom, L.W., 2009. Thelytoky in Hymenoptera with Venturia canescens and Leptopilina clavipes as case studies, in: Lost Sex. Springer, pp. 347–375.

Maynard-Smith, J., 1978. The evolution of sex. Cambridge University Press Cambridge.

McConnell, M.W., Fitzpatrick, M.J., 2017. "Foraging" for a place to lay eggs: A genetic link between foraging behaviour and oviposition preferences. PLoS One 12, e0179362. https://doi.org/10.1371/journal.pone.0179362

Normark, B.B., 2003. The evolution of alternative genetic systems in insects. Annual review of entomology 48, 397–423.

Osborn, K.A., Robichon, A., Burgess, E., Butland, S., Shaw, R.A., Coulthard, A., Pereira, H.S., Greenspan, R.J., Sokolowski, M.B., 1997. Natural Behaviour Polymorphism Due to a cGMP-Dependent Protein Kinase of Drosophila. Science 277, 834–836. https://doi.org/10.1126/science.277.5327.834

Osborne, K., Robichon, A., Burgess, E., Butland, S., Shaw, R., Coulthard, A., Pereira, H., Greenspan, R., Sokolowski, M., 1997. Natural behavior polymorphism due to a cGMP-dependent protein kinase of Drosophila. Science 277, 834–836.

Otto, S.P., 2009. The evolutionary enigma of sex. the american naturalist 174, S1–S14.

Pearce, L.R., Komander, D., Alessi, D.R., 2010. The nuts and bolts of AGC protein kinases. Nature reviews Molecular cell biology 11, 9–22.

Pelosse, P., Bernstein, C., Desouhant, E., 2007. Differential energy allocation as an adaptation to different habitats in the parasitic wasp Venturia canescens. Evolutionary Ecology 21, 669–685. https://doi.org/10.1007/s10682-006-9145-5

Pelosse, P., Jervis, M.A., Bernstein, C., Desouhant, E., 2011. Does synovigeny confer reproductive plasticity upon a parasitoid wasp that is faced with variability in habitat richness? Biological Journal of the Linnean Society 104, 621–632. https://doi.org/10.1111/j.1095-8312.2011.01741.x

Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R., 2017. Evolutionary history of the Hymenoptera. Current Biology 27, 1013–1018.

R Core Team, 2017. R: A language and environment for statistical computing.

Reaume, C.J., Sokolowski, M.B., 2009. cGMP-dependent protein kinase as a modifier of behaviour. cGMP: generators, effectors and therapeutic implications 423–443.

Rittschof, C.C., Robinson, G.E., 2016. Chapter Five - Behavioral Genetic Toolkits: Toward the Evolutionary Origins of Complex Phenotypes, in: Orgogozo, V. (Ed.), Current Topics in Developmental Biology. Academic Press, pp. 157–204. https://doi.org/10.1016/bs.ctdb.2016.04.001

Rogers, D., 1972. The Ichneumon Wasp Venturia Canescens: Oviposition and Avoidance of Superparasitism 15.

Sacomoto, G.A., Kielbassa, J., Chikhi, R., Uricaru, R., Antoniou, P., Sagot, M.-F., Peterlongo, P., Lacroix, V., 2012. KISSPLICE: de-novo calling alternative splicing events from RNA-seq data. BMC Bioinformatics 13, S5. https://doi.org/10.1186/1471-2105-13-S6-S5

Salt, G., 1976. The hosts of Nemeritis canescens a problem in the host specificity of insect parasitoids. Ecological Entomology 1, 63–67. https://doi.org/10.1111/j.1365-2311.1976.tb01205.x

Schlossmann, J., Desch, M., 2009. cGK substrates. cGMP: Generators, Effectors and Therapeutic Implications 163–193.

Schneider, M. V., Beukeboom, L.W., Driessen, G., Lapchin, L., Bernstein, C., Van Alphen, J.J.M., 2002. Geographical distribution and genetic relatedness of sympatrical thelytokous and arrhenotokous populations of the parasitoid venturia canescens (Hymenoptera). Journal of Evolutionary Biology 15, 191–200. https://doi.org/10.1046/j.1420-9101.2002.00394.x

Sokolowski, M.B., 1980. Foraging strategies of Drosophila melanogaster: A chromosomal analysis. Behavior genetics 10, 291–302.

Sokolowski, M.B., Pereira, H.S., Hughes, K., 1997. Evolution of foraging behavior in Drosophila by density-dependent selection. Proceedings of the National Academy of Sciences 94, 7373–7377.

Stouthamer, R., Luck, R.F., Hamilton, W., 1990. Antibiotics cause parthenogenetic Trichogramma (Hymenoptera/Trichogrammatidae) to revert to sex. Proceedings of the National Academy of Sciences 87, 2424–2427.

Struk, A.A., Mugon, J., Huston, A., Scholer, A.A., Stadler, G., Higgins, E.T., Sokolowski, M.B., Danckert, J., 2019. Self-regulation and the foraging gene (PRKG1) in humans. Proceedings of the National Academy of Sciences 116, 4434–4439.

Tobback, J., Mommaerts, V., Vandersmissen, H.P., Smagghe, G., Huybrechts, R., 2011. Age-and task-dependent foraging gene expression in the bumblebee Bombus terrestris. Archives of insect biochemistry and physiology 76, 30–42.

van der Kooi, C.J., Matthey-Doret, C., Schwander, T., 2017. Evolution and comparative ecology of parthenogenesis in haplodiploid arthropods. Evol Lett 1, 304–316. https://doi.org/10.1002/evl3.30

Wenseleers, T., Billen, J., Arckens, L., Tobback, J., Huybrechts, R., Heylen, K., Gobin, B., 2008. Cloning and expression of PKG, a candidate foraging regulating gene in Vespula vulgaris. Animal Biology 58, 341–351.

West, S., Flanagan, K., Godfray, H., 1996. The relationship between parasitoid size and fitness in the field, a study of Achrysocharoides zwoelferi (Hymenoptera: Eulophidae). Journal of Animal Ecology 631–639.

Whitfield, C.W., Cziko, A.M., Robinson, G.E., 2003. Gene expression profiles in the brain predict behavior in individual honey bees. Science 302, 296–299. https://doi.org/10.1126/science.1086807

Accessibility and Benefit-Sharing section

Data Accessibility statement

Data were deposited in the GEO repositories from NCBI database with the accession code GSE194171.

Benefit-Sharing Statement

Benefits from this research accrue from the sharing of our data and results on public databases as described above.

Author Contributions

A.G., E.D., L.M., C.V.H. and I.A. designed research; A.G. and D.L. performed the research; A.G., V.L., D.L. and A.E.F. analyzed data; A.G., E.D., L.M., C.V.H. and I.A. wrote the paper.

Tables and Figures

Hosted file

`image1.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure 1 - The *for* gene in *V. canescens.* A) Maximum likelihood phylogenetic tree of the *for* gene in insects. Sequences used from 39 species of insects: hemiptera (2), orthoptera (2), coleoptera (4), lepidoptera (6), diptera (15), hymenoptera (10). Sequence from *V. canescens* figured in bold within hymenoptera cluster.

One crustacean (*Daphnia pulex*) and one mammal (*Mus musculus*) sequences were used as outgroups. Phylogeny was constructed from the 533 aminoacids located in the conserved C-terminal region of the protein. Branch support was estimated by likelihood-ration test, aLRT values [?]0.90 were indicated with empty circles, and [?]0.95 with full circles. B) Schematic of the *for* gene model and 17 associated isoforms identified in *V. canescens.* The *for* gene located on the scaffold 64 of *V. canescens* genome was constituted of 13 exons (dark blue boxes). The two conserved functional protein domains (two cGMP binding domains and protein kinase domain) were located at the 3' extremity of the gene sequence and were represented under the genome sequence (in blue and red respectively). The 5' part of the gene exhibited variations, with 4 distinct promotors identified (*pr1-4* , dark blue arrows), and the corresponding transcripts were labelled accordingly (*Vcan27706* to *Vcan27709*). A total of 17 isoforms were identified within RNA-seq libraries issued from sexual and asexual *V. canescens* heads, and were all represented under the genome sequence with coding sequences figured in dark grey boxes and non-coding sequences (5' and 3' UTRs) in pale grey boxes. The amplicon used to quantify *for* transcript with RT-qPCR was common to almost all isoforms (excepted *Vcan_27707i12*) and figured in green.

Table 1 – Decreased polymorphism at the *for* locus in asexual population. Characterisation of *for* SNPs that differed in frequency across sexual and asexual populations: transcript column indicated the isoform cluster containing the variant; the region and position column indicated if the SNP is located within the coding sequence (CDS) or outside (UTR) and its position considering the longest transcript sequence figured into bracket; polymorphism indicated the different bases identified at the SNP position; sexual and asexual population columns contained the number of reads corresponding to each variant in brackets. In the last two columns information related to the type of mutation (synonymous or non-synonymous, when SNPs occurred within CDS), and aminoacid change (only for the non-synonymous mutation) are informed.

Hosted file

`image2.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Hosted file

`image3.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure 2 - Transcriptome comparison between sexual and asexual population. A) Two-dimensional principal component analysis based on head transcriptomes from sexual (black) and asexual (red) populations. Samples, based on the expression profile of 14,106 genes, clustered according to population. B) Volcano plot illustrated variations in gene expression between sexual and asexual populations, the y axis measured the statistical significance (FDR) while the x axis indicated the magnitude of change. Grey points represented the 1,539 DEG, while genes not differentially expressed were represented by blue dots including the*for* gene pointed in red. C) Comparison of normalized counts number of the *for* gene between the 3 asexual (red) and 3 sexual (black) RNA-seq libraries.

Table 2 – Behavioural genes belonging to the DEGs between sexual and asexual populations. Among the 1539 DEGs between sexual and asexual populations, 26 transcripts were annotated has related to behaviour GO term, or one of its child GO terms. The population in which the transcript was overexpressed was indicated as well as its fold change (log2), with positive values when overexpressed in sexual population, and negative values in asexual population.

Hosted file

`image4.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Hosted file

`image5.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure 3 - Foraging behaviour differences between sexual and asexual populations of *V. canescens*. A) Total time spent on host patches (PRT) according to the population. B) Number of cockings, corresponding to the number of eggs deposited into hosts according to the population. C) Number of switches between the 2 host patches according to the population. Each population consisted of 17 females from asexual or sexual population.

Hosted file

`image6.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure 4 – Comparison of *for* expression between sexual and asexual populations, in relation with fertility. A) Variations in the *for* gene expression between sexual and asexual populations. B) Individual *for* expression according to egg load in sexual (black) or asexual (red) females. Black and red lines corresponded to the values predicted by the model in sexuals and asexuals, respectively. C) Individual *for* expression according to the number of eggs laid in asexual females. Circles represent data measured in females: 17 asexuals (red) and 17 sexuals (black). Expression of the gene *for*is indicated relatively to the median individual.

Appendix - Supplementary data

Table S1 – Set of *for* (PKG protein) ortholog sequences used in the phylogenetic analysis.

Table S2 - RNA-seq librairies information.

Table S3 - Primers sequences used for qPCR.

Table S1- Raw results of qPCR

Figure S1 – Information about quality of RT-qPCR results.

Figure S2 – Comparison of the 4 *for* main isoforms across sexual and asexual population.

Table S1- PKG1 sequences used in the phylogenetic analysis.

species	ordre	sequence ID (ensembl)	sequence ID (uniprot)
Acyrthosiphon pisum	hemiptera	ACYPI008877	J9K7J2
Rhodnius prolixus	hemiptera	RPRC000321-RA	T1H8H9
Aedes aegypti	diptera	AAEL007826	Q170R4
Anopheles gambiae	diptera	AGAP008863	A0A1S4H065
Belgica antartica	diptera	IU25_00750	
Culex quinquefasciatus	diptera	CPIJ005276	
Culex quinquefasciatus	diptera	CPIJ005277	
Drosophila ananassae	diptera	FBtr0119359	B3MP45
Drosophila erecta	diptera	FBtr0408846	B3N351
Drosophila grimshawi	diptera	FBtr0148601	B4JQH5
Drosophila persimilis	diptera	FBtr0185081	B4G9E3
Drosophila pseudoobscura	diptera	FBtr0379386	Q29MD8
Drosophila simulans	diptera	FBtr0222650	B4Q9N2
Drosophila yakuba	diptera	FBtr0261343	B4NXZ5
Drosophila melanogaster	diptera	FBpp0088350	
Lucilia cuprina	diptera	KNC24599	A0A0L0BZC8
Mayetiola destructor	diptera	Mdes008791-RA	
Teleopsis dalmanni	diptera	TDAL000635-RA	
Teleopsis dalmanni	diptera	TDAL011481-RA	
Anoplophora glabripennis	coleoptera	AGLA020541	
Dendroctonus ponderosae	coleoptera	ENN75590	N6U233
Tribolium castaneum	coleoptera	TC034109_001	A0A139WDE4
Diabrotica virgifera virgifera	coleoptera	ABI97017.1	Q06BQ9
Apis mellifera	hymenoptera	GB49908	A0A088AEM0
Atta cephalotes	hymenoptera	XM_012208338.1	A0A158P1S0
Bombus impatiens	hymenoptera	BIMP24832	
Bombus terrestris	hymenoptera	XM_003400304.3	
Nasonia vitripenis	hymenoptera	NV12180-RA	K7ISX7
Solenopsis invicta	hymenoptera	XM_011170343.1	A0A097KJ21
Pheidole pallidula	hymenoptera	EF999975	ABW22623.1
Pogomyrmex barbatus	hymenoptera	AY800387.1	XP_011637860.1
Vespula vulgaris	hymenoptera	EF136648.1	ABL74445.1
Venturia canescens	hymenoptera	TCONS_00027709	
Danaus plexipus	lepidoptera	DPOGS208453	
Heliconius melpomene	lepidoptera	HMEL017894-RA	
Melitaea cinxia	lepidoptera	MCINX011043-RA	
Bombyx mori	lepidoptera	AF465600.1	Q8ST45
Lobesia botrana	lepidoptera	DQ666642.1	ABG56236.1
Mythimna separata	lepidoptera	GQ844298.1	ACX46913.1
Schistocerca gregaria	orthoptera	ADP94162.2	
Locusta migratoria	orthoptera	FJ214984	ADP94162.2
Daphnia pulex	crustacea	EFX84317	E9G886
Mus musculus	vertebrate	XP_006526831.1	

Table S2 - RNA-seq librairies information.

ID	population	number of paired reads	alignment rate
AMXC-1	sexual	14,556,391	94.12%
AMXC-2	sexual	14,883,970	92.30%
AMXC-3	sexual	18,465,742	93.32%
AMXC-4	asexual	12,999,549	90.76%
AMXC-5	asexual	15,146,936	89.97%
AMXC-6	asexual	14,967,273	91.21%

Table S3 - Primers sequences used for RT-qPCR.

Gene	Description	Forward primer sequence	Reverse primer sequence
Rpl32	Ribosomal protein L 32	5'-GCGTTTCAAGGGTCAGTTCT-3'	5'-AGCGATCTCTGCAC
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase	5'-TGTGTCCGGATGTACCTGAGT-3'	5'-TTAAATACACGCTC
For	Foraging gene	5'-GGGTCGAACTCGTCCAAATA-3'	5'-CTTGGCCATCGAGC

Table S4 - Raw results of qPCR

female_id	*For* (C_q)	*Rpl32* (C_q)	*Gapdh* (C_q)
26	28	27.19	28.05
38	27.8	25.74	26.41
4	30.19	27.76	28.16
16	31.42	27.45	28.29
60	30.44	27.25	27.12
53	31.8	27	27.65
5	27.13	25.43	26.52
14	30.94	27.76	27.42
52	27.77	25.84	26.74
40	28.18	25.66	27.64
32	31.26	26.78	27.62
50	29.27	27.52	28.31
2	31.9	28.21	28.55
19	31.13	28.34	29.21
55	28.13	27.18	28.23
59	25.97	25.96	26.89
7	31.21	29.12	29.95
48	29.95	28.34	29.29
#12	34.96	33.63	33.85
28	31.25	27.58	28.29
57	26.99	27.59	29.04
9	29.11	26.59	27.23
37	28.45	26.67	27.64
31	29.67	27.43	27.37
58	30.32	26.96	27.49
1	29.72	26.03	26.5
49	28.47	26.5	27.06
15	29.1	26.58	27.21
8	28.57	28.45	29.65
54	27.68	25.42	27.08
3	31.06	27.46	27.72
6	27.56	27.65	28.33
41	30.22	27.5	27.9
39	29.68	27.76	28.37
18	29.84	28.47	28.56
23	30.77	27.53	27.86
29	29.18	27.33	27.7
27	31.06	28.05	28.01
21	26.84	26.22	27.1
13	26.61	26.06	27.27
61	26.64	26.15	27.73
45	26.4	26.42	27.52

Hosted file

`image8.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure S1 – Information about quality of RT-qPCR results. Boxplot of quantification cycle (C_q) measured on 34 individual heads for 3 genes: *for* , *rpl32* , *gapdh* . The *rpl32* and *gapdh* used as reference gene

to normalised*for* expression exhibited less variation among samples than the*for* gene (F-test; p-val<0.05) but do not differ between them. C_q values were lower in reference genes than in*for* gene, indicating that *rpl32* and *gapdh* have a higher expression compared to *for* gene.

Hosted file

`image9.emf` available at https://authorea.com/users/504994/articles/584220-the-for-gene-as-one-of-the-drivers-of-foraging-variations-in-a-parasitic-wasp

Figure S2 - Comparison of the 4 *for* main isoforms across sexual and asexual population. None of the 4 *for* transcripts (*Vcan27706, Vcan27707, Vcan27708, Vcan27709*) exhibited significant differential expression across asexual and sexual populations (FDR>0.05).