

***De novo* transcriptome sequencing of the northern fowl mite, *Ornithonyssus sylviarum*, shed light into parasitiform poultry mites evolution and its chemoreceptor repertoires**

5 Biswajit Bhowmick¹, Huaqing Chen¹, Jesus Lozano-Fernandez², Joel Vizueta³,
Rickard Ignell⁴, and Qian Han^{1*}

¹Laboratory of Tropical Veterinary Medicine and Vector Biology, School of Life and
Pharmaceutical Sciences, Hainan University, Haikou, Hainan 570228, China.

10 ² Institute of Evolutionary Biology (CSIC-UPF), Pg. Marítim de la Barceloneta 39-47,
08003 Barcelona, Spain

³ Villum Centre for Biodiversity Genomics, Section for Ecology and Evolution,
Department of Biology, University of Copenhagen, Universitetsparken 15, 2100
Copenhagen, Denmark

15 ⁴ Disease Vector Group, Unit of Chemical Ecology, Department of Plant Protection
Biology, Swedish University of Agricultural Sciences, 23053, Alnarp, Sweden

*Correspondence: Qian Han Email: qianhan@hainanu.edu.cn

Abstract

20 The poultry red mite (PRM), *Dermanyssus gallinae* (De Geer), and the northern fowl
mite (NFM), *Ornithonyssus sylviarum* (Canestrini and Fanzago), are the most serious
pests of poultry, both of which have an expanding global prevalence. Research on
NFM has been constrained by a lack of genomic and transcriptomic data. Here, we
report and analyze the first transcriptome data for this species. A total of 28,999
25 unigenes were assembled, of which 19,750 (68.10%) were annotated using seven
functional databases. The biological function of these unigenes was predicted using
the GO, KOG, and KEGG databases. To gain insight into the chemosensory
receptor-based system of parasitiform mites, we furthermore assessed the gene
repertoire of gustatory receptors (GRs) and ionotropic receptors (IRs), both of which
30 encode putative ligand-gated ion channel proteins. While these receptors are well
characterized in insect model species, our understanding of chemosensory detection
in mites and ticks is in its infancy. To address this paucity of data, we identified 9
IR/iGluRs and 2 GRs genes by analyzing transcriptome data in the NFM, while 9
GRs and 41 IR/iGluRs genes were annotated in the PRM genome. Taken together,
35 the transcriptomic and genomic annotation of these two species provide a valuable
reference for studies of parasitiform mites, and also helps to understand how
chemosensory gene family expansion/contraction events may have been reshaped by
an obligate parasitic lifestyle compared with their free-living closest relatives. Future
studies should include additional species to validate this observation, and functional
40 characterization of the identified proteins as a step forward in identifying tools for
controlling these poultry pests.

Key words: ectoparasites, *de novo* assembly, genomics, chemosensory receptors,

phylogeny, Parasitiformes.

45 **Introduction**

The northern fowl mite (NFM), *Ornithonyssus sylviarum*, and the poultry red mite (PRM), *Dermanyssus gallinae* (Parasitiformes: Mesostigmata: Dermanyssoidea), are small (< 1 mm) obligate blood-feeding ectoparasites of birds that cause significant economic losses to the poultry industry, and a negative impact on animal welfare. The entire life-cycle of NFM occurs on the host, while PRM lives off-host in the cracks and crevices in poultry houses (Tomley et al., 2018; Sparagano and Ho, 2020; Murillo et al., 2020). The development of novel intervention strategies to control poultry mite population is of great interest. In this context, genomic information is essential, but not available for the NFM. Until now, only a few examples of nucleotide sequences referring to the mitochondrial genome of NFM are deposited in the NCBI database (Roy et al., 2009; Bhowmick et al., 2019). With the development of high throughput sequencing (HTS) technologies, scientists have been able to study the transcriptomes of Acari (Parasitiformes and Acariformes orders), with a few reference genomes currently available. The first transcriptome analysis was carried out on PRM (Schicht et al., 2014), for which the draft genome assembly was reported based on combined PacBio and Oxford Nanopore MinION long-read *de novo* sequencing (Burgess et al., 2018).

Similar to most insects (Joseph and Carlson, 2015), non-insect arthropods, such as parasitiform ectoparasites (*Ixodes scapularis*, *O. sylviarum*, *D. gallinae*, and *Varroa destructor*), rely heavily on their sense of smell to discriminate and locate resources, including their vertebrate hosts (Sonenshine, 2004; Gay et al., 2020; Light et al., 2020; Faraone et al., 2020; Josek et al., 2021). To do so, non-insect arthropods seem to rely on chemosensory receptors from two gene families: gustatory receptors (GRs) and ionotropic receptors (IRs) (Josek et al., 2018). In insects, the GR family of seven-transmembrane (TM) proteins comprises up to hundreds of highly divergent (only 8–12% sequence identity) receptors, that are involved in the detection of various tastants, including e.g., bitters, sugars and cuticular pheromone compounds, as well as carbon dioxide (Vosshall et al., 1999; Clyne et al., 2000; Chyb et al., 2003; Jones et al., 2007). The IRs are members of the second largest family of chemoreceptors in insects, and represents a highly divergent subfamily of ionotropic glutamate receptors (iGluRs), which include the NMDA-type receptors, AMPA-type receptors, and kainate-type receptors that have ultra-conserved roles in synaptic transmission in animals, where the latter two are grouped together as non-NMDA types (Benton et al., 2009; Sobolevsky et al., 2009). The IRs involved in the detection of chemical stimulus (e.g. odorants), require a co-receptor (IR8a, IR25a or IR76b) to function (Abuin et al., 2011). These co-receptors along with IRs are involved in different physiological functions, such as the perception of temperature or humidity and are conserved across insects and other arthropods (Enjin et al., 2016; Ni et al., 2016).

In the present study, we performed a high-coverage sequencing and *de novo* assembly of the parasitiform mite *O. sylviarum*. Based on this assembly, we discuss functional categories and annotations, functional genes associated with various physiological functions, and transcriptional factors of relevant genes and gene families. Gene ontology (GO) and KEGG enrichment analyses were conducted to identify the biological functions and predict the pathways of key genes. Our whole-transcriptome data provides important molecular information on mite biology and might provide new opportunities for the development of novel control approaches against mite infestation, such as the identification of new drug targets or the development of potential vaccine candidates. The generation of the HTS data, furthermore, allowed for the investigation of chemosensory gene family evolution, such as IRs and GRs across arthropods (Eyun et al., 2017; Vizueta et al., 2020a). In both mite species, we identified and classified the putative IR/iGluR and GR unigenes using HTS data. These genes were analyzed using a series of bioinformatics tools to exhaustively understand their characteristics, including their protein structure, molecular domain architecture, and phylogenetic relationships with IR/iGluRs of other mite and tick species. In addition, we performed an automated refining and validation of the candidate gene model via ORCAE (Online Resource for Community Annotation of Eukaryotes), which offers a community-based genome annotation platform in which multiple researchers from around the world can contribute new information on gene structure, function, and gene expression atlas (Sterck et al., 2012; Yssel et al., 2019; Burgess et al., 2019). This study yields complementary information on chemosensory receptor gene families in poultry mites and other mite and tick species and provides an enriched set of gene data of traditional iGluRs, conserved IRs, and GRs, which are essential for further research on gene evolution in chemoreceptor families.

Materials and methods

Mite collection, RNA extraction and quality determination

The NFM s were collected from a commercial layer poultry farm, as previously described (Wenchang, Hainan, China; 19° 29' 1" North, 110° 46' 18" East) (Bhowmick et al., 2019). All life-stages and sexes of *O. sylviarum* mites were prepared for total RNA extraction. The samples were washed in ice-cold phosphate buffer saline (PBS) in order to clean the surfaces of the mites and then stored at -80°C for later use. A pooled sample with multiple individuals was homogenized by snap freezing in liquid nitrogen, and then RNA extraction was performed according to the TRIzol reagent protocol (Invitrogen, CA, USA). RNA parameters and quality were assessed using an Agilent BioAnalyzer, and the concentration and purity were assessed using the NanoDrop 2000.

cDNA library preparation and next generation sequencing

125 The pooled RNA sample was sent to BGI (Beijing Genomics Institute, China)
for library generation and next generation sequencing (NGS). Briefly, the mRNA was
isolated using the oligo (dT)-attached magnetic beads, and then a fragmentation
buffer was used to break the mRNA into short fragments as templates for first-strand
130 cDNA synthesis. The resulting first-stranded cDNA was generated using reverse
transcriptase with random hexamer primer mix, and then the resulting
second-stranded cDNA was synthesized using DNA polymerase I, buffer, dNTPs and
RNase H. Subsequently, the synthesized double-stranded cDNA was purified and
blunt end repair was carried out, adding a base A and a linker to the 3' end. After the
end repair and ligation of adaptors, the product was amplified by PCR and purified
135 using a PCR purification kit to create a single-stranded circular DNA library. Finally,
a sequencing library was performed using the BGISEQ-500 platforms.

***De novo* transcriptome assembly, annotation, and functional classification**

Prior to assembly and mapping, clean reads were obtained by removing those
with low quality, adapter contamination, and poly-N from the raw reads. Meanwhile,
140 phred score, Q20 (%) , Q30 (%) and GC content in the clean reads were calculated
(Cock et al., 2010). Trinity (v2.0.6) package was used to assemble the high-quality
clean data, and the assembled unigenes were then analyzed using the following
databases: Nr (<ftp://ftp.ncbi.nlm.nih.gov/blast/db>; non-redundant protein sequences),
Nt (<ftp://ftp.ncbi.nlm.nih.gov/blast/db>; non-redundant nucleotide sequences), Pfam
145 (<http://pfam.xfam.org>; protein family), KOG / COG
(<https://www.ncbi.nlm.nih.gov/COG/>; clusters of orthologous groups), Swiss-Prot (a
manually annotated and reviewed protein sequence database), KEGG
(<http://www.genome.jp/kegg>; Kyoto Encyclopedia of Genes and Genomes), GO
(<http://geneontology.org>; Gene Ontology), and CD-HIT
150 (<https://gvolante.riken.jp/analysis.html>). In addition, TransDecoder
(<https://transdecoder.github.io>) software was then used to determine coding regions
(CDS) in unigenes. To identify genes encoding transcription factor (TF) families, all
unigenes were analyzed using the Animal Transcription Factor Database
(AnimalTFDB), and protein domains were identified using HMMER (Finn et al.,
155 2015). Finally, to assess the completeness of the transcriptomes, we evaluated its
content by benchmarking it against the arthropoda single-copy orthologous database
using BUSCO V3 (Simão et al., 2015).

Identification, annotation and phylogenetic analyses of IR/iGluRs and GRs

160 The IR/iGluR and GR gene family members were annotated in the genome of
the PRM (BioProject accession number PRJNA487003; Burgess et al., 2018). Briefly,
the BITACORA pipeline (Vizueta et al., 2020b) was used to conduct homology-based
searches and annotate the IR/iGluR and GR genes encoded in the input genome, using
165 a curated database of these genes enriched in chelicerate species. The output of
BITACORA included gene-finding format (GFF), with both curated and newly
identified gene models, and a FASTA file of predicted protein sequences (Vizueta et al.,

2018). To identify the surveyed receptors in the NFM transcriptome, we then used a three-step approach to annotate the IR/iGluRs and GRs in the assembled transcripts.

170 Firstly, IR/iGluR and GR unigenes from the transcriptomes were retrieved by the functional annotation results based on seven functional databases (NR, KEGG, KOG, NT, SwissProt, Pfam, and GO). Then, iterative tBLASTn and PSI-BLAST searches were performed with previously identified amino acid sequences of the parasitiform tick *I. scapularis* and the parasitiform mites *D. gallinae*, *Tropilaelaps mercedesae* and

175 *V. destructor* as initial queries in order to identify additional transcripts (Altschul et al., 1997). All of the candidate genes were manually verified by BLASTx against the non-redundant (NR) database at NCBI, and the open reading frames (ORFs) were identified using the ORFfinder at NCBI. We also tested signature patterns of Pfam domain architectures for IR/iGluRs (PF01094; PF00060; PF10613) and GRs

180 (PF08395 and PF06151) using the HMMER software (Potter, et al., 2018). Thereafter, the amino acid sequences were blasted against the pfam (<http://pfam.xfam.org/>) database to obtain unigene annotation information (Finn et al., 2015). Lastly, prediction of membrane protein topology and signal peptides were performed with TOPCONS and TMHMM (Bernsel, et al. 2009). Amino acid sequences of IR/iGluRs

185 predicted in both *O. sylviarum* and *D. gallinae*, along with *T. mercedesae*, *I. scapularis*, *Galendromus* (= *Typhlodromus*, = *Metaseiulus*) *occidentalis*, *V. destructor*, the vinegar fly *D. melanogaster*, and the crustacean *Daphnia magna* were aligned using MAFFT default settings. The GR sequences from *O. sylviarum* and *D. gallinae* were analyzed together with those from *T. mercedesae*, *I. scapularis*, *M. occidentalis*, *V. destructor*,

190 *Tetranychus urticae*, *D. melanogaster*, and *Daphnia pulex*. The amino acid sequences used for phylogenetic tree construction are listed in Supplementary Material (Supplementary Table S1 and S2; Vizueta et al., 2018). A phylogenetic tree was performed with PhyML (maximum likelihood) software using the best-fitting likelihood model LG+G+I+F as calculated by SMS (Smart Model Selection) (Lefort et al., 2017; Katoh et al., 2019). A phylogenetic tree branch support was calculated using a new, fast, and approximate maximum likelihood-ratio test (aLRT), and the tree was visualized in EvolView (He et al. 2017).

***Ab initio* modelling of protein sequence**

200 The three-dimensional protein structure and function of selected proteins of *O. sylviarum* were predicted using an online version of I-TASSER (Iterative Threading Assembly Refinement) (Roy et al., 2010). The crystal structure of the AMPA subtype ionotropic glutamate receptor (PDB code: 3KG2) served as a template for the Osyl_iGluR7. For the Osyl_IR9, the crystal structure of *Rattus norvegicus* (PDB code: 5kuf) served as a template. Pymol was used to visualize the complex model. The

205 the quality of the resulting structural homology model was evaluated in comparison with the quality of the template.

Results and discussion

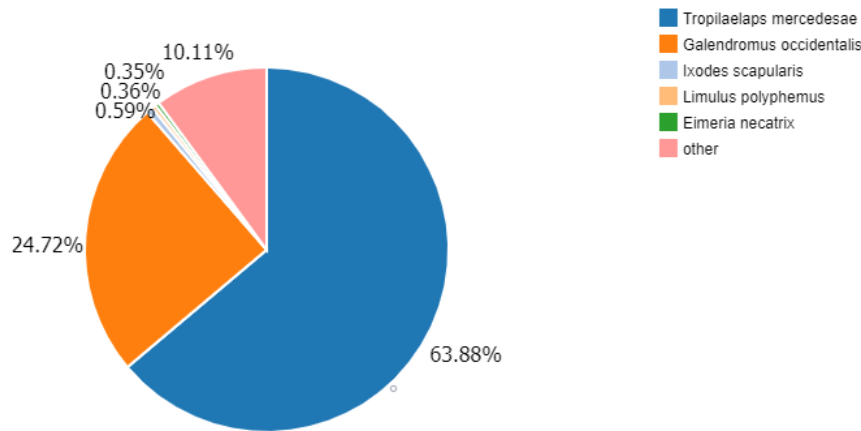
***De novo* assembly, functional annotation and classification of the *O. sylviarum* transcriptome**

210

To gain relevant molecular information on *O. sylviarum*, as a mean to identify novel targets for pest control, we used the BGISEQ-500 next generation sequencing platform, which provided a total of 143.74 million raw reads (totaling roughly 19.72 GB of sequence data), of which 91.43 % were clean reads. Among the 131.41 million
215 clean reads, 96.98% and 88.75 % had quality scores above the Q20 and Q30 levels, respectively (Supplementary Table S3). After Trinity *de novo* assembly of transcriptome, in total 28,999 unigenes were obtained with an average length of 1,635 bp. Total sequencing length, N50 and GC content were 47,424, 969 bp, 2,907 bp and 48.33%, respectively (Supplementary Table S4; Supplementary Figure
220 S1). The extent of completeness and duplication of the *O. sylviarum* gene set were assessed using the arthropod-specific Benchmarking Universal Single-Copy Ortholog (BUSCO) genes (Simão et al., 2015). We retrieved 90.53% complete orthologs (C), being 24% duplicated (D) only 2.53% fragmented orthologs (F), and 6.94% missing (M) from the assembly (Supplementary Figure S2). BUSCO indicated a high degree
225 of completeness at 93.06%, which suggests that most of our transcriptome assembly is significantly complete in terms of gene repertoire (Supplementary Figure S2). In total 19,750 unigenes were generated by the seven functional databases: 17,783 (61.32%) in NR, 6,353 (21.91%) in Nt, 12,956 (44.68%) in KOG, 13,528 (46.65%) in Swiss-Prot, 14,840 (51.17%) in Pfam, 5,853 (20.18%) in GO and 13,661 (47.11%)
230 in KEGG annotation. The remaining 9,249 (31.89%) unigenes were not annotated, suggesting that some of them may represent lineage-specific novel genes, in addition to the typical presence of assembly artifacts or non-coding RNA. Similarly, 27,547 unigenes (67.86%) of the mite *D. gallinae* were classified as unknown function, highlighting the need for functional genomic studies for poultry mite species (Huang
235 et al., 2020).

Based on the unigenes with GO annotation (20.18%), these unigenes were primarily divided into three separate ontologies: biological processes, molecular functions and cellular components. The most abundant categories were classified into
240 catalytic activity (2,667 unigenes), binding (2,620 unigenes) and cellular processes (1,950 unigenes) (Supplementary Figure S3). Apart from the GO ontology analysis, a total of 13,661 (47.11%) unigenes were used for KEGG analysis, which identified six main pathway branches (Supplementary Table S5). Among these, ‘metabolism’, ‘environmental information processing’ and ‘cellular process’ were the most
245 dominant categories, in which metabolism related pathways occupied the bulk part in NFM. These results can provide a valuable reference for investigating specific biological, molecular functions and pathway analyses of *O. sylviarum* transcriptome (Supplementary Figure S4). According to the NR database, the top-hit species was the honey bee parasitic mite *T. mercedesae* (63.88%), followed by the predatory mite
250 *G. occidentalis* (24.72%) and the deer tick *I. scapularis* (0.59%), suggesting that the NFM transcriptome matches well with other Parasitiformes mites. Thus, poultry

mites share most of the orthologous genes with the parasitiform mites *T. mercedesae* and *G. occidentalis* rather than with the parasitiform tick *I. scapularis* (Figure 1), which is in agreement with a previous study (Huang et al., 2020). The KOG analyses can provide useful information of the annotated unigenes based on gene-based definition of orthology. In total of 44.68% unigenes with non-redundant database hits were grouped into 25 KOG categories (Supplementary Figure S5). For *O. sylviarum*, the largest categories were ‘general function prediction only’, followed by ‘signal transduction mechanisms’, ‘function unknown’, and ‘posttranslational modification, protein turnover’. The smallest unit of classification were ‘cell motility’, ‘nuclear structure’ and ‘nucleotide transport and metabolism’. The raw reads have been deposited in the Sequence Read Archive (SRA) database under the accession number PRJNA675700, while the Transcriptome Shotgun Assembly (TSA) has been submitted with accession number GIXZ00000000.



265

Figure 1: The species distribution of BLASTx matches of the transcriptome unigenes. Each section shows the number of top BLASTx hits.

Loss or reduction of highly conserved transcription factors in *O. sylviarum*

Transcription factors (TFs) are key proteins that bind specifically to DNA, and are an integral part of gene expression at the transcriptional level. Currently, there is little information available for the structures and biological functions of TFs in mites. In the present study, 2,191 putative unigenes were identified and found to be distributed across 65 TF families. Among these, zf-C2H2 was the most abundant family (668 unigenes), followed by Homeobox (172 unigenes), and bHLH (125 unigenes) (Figure 2). Homeobox genes were present in a range of arachnids that contain a highly conserved homeodomain. Widespread retention and pervasive gene duplication of homeobox gene has been observed in Arachnida class (e.g., spiders, ticks, scorpions, and mites), suggesting implications for the field of evolutionary biology (Leite et al., 2018). Genomic analyses have also reported that mite species exhibit dynamic rearrangements of homeobox clusters, along with loss of highly

280

conserved TFs (Hoy, 2009; Leite et al., 2018; Grbic et al., 2011). The basic helix-loop-helix (bHLH) is one of the largest family of transcriptional regulatory proteins, and are present in organisms from yeast to humans, where their members are participating in regulating a wide range of biological and developmental processes, such as the development of the nervous system and muscles, as well as responding to environmental factors (Jones, 2004). Notably, a number of specific TFs that are highly conserved in most arthropods were not detected in the *O. sylviarum* transcriptome, including circadian rhythm (so-called Clock-Cycle), E78, HR3, Har-AP-4, HR39, EcR, Abd-A, ERR, Ato, FTZ-F1, ATF-3, HR96, Dip3, NRs HNF4, HR78, CREB, HR83, and Dnato3 (Guo et al., 2018). The absence of these TFs is highly intriguing, and clearly demands further investigation, using different transcriptome or genomic data in order to validate this result.

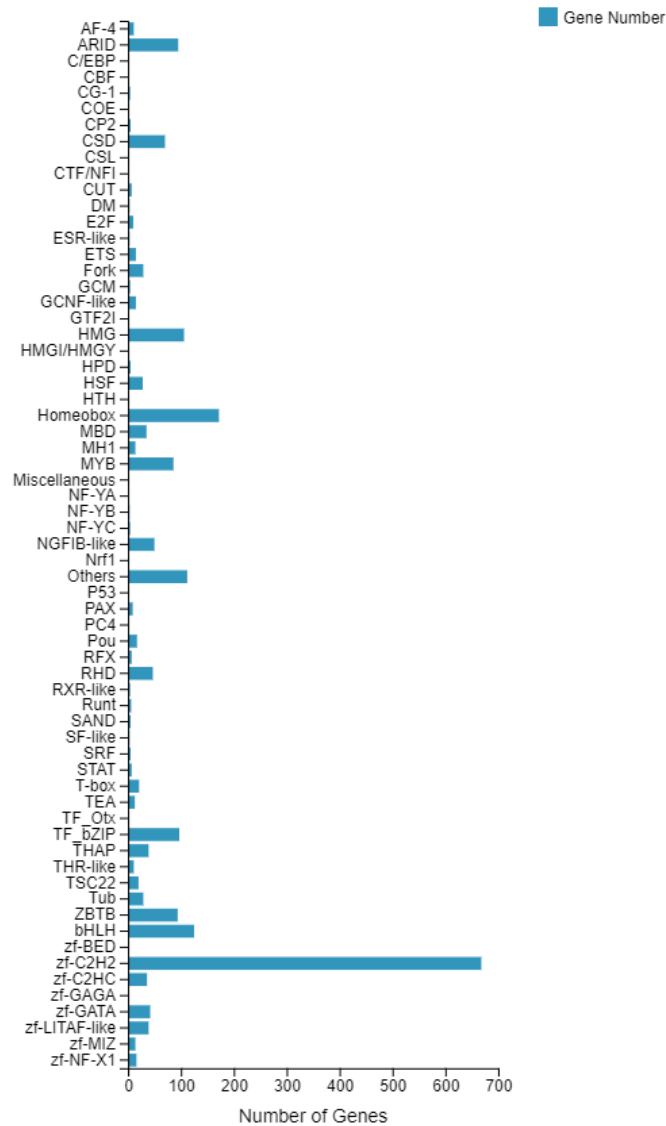


Figure 2: Family distribution of the transcription factors (TFs) in the NFM transcriptome. The number of genes identified in each TF family are represented in X-axis.

300 Identification and interactive web-based annotation of chemosensory receptors

Annotation of the IRs and GRs by a bioinformatics-based approach, coupled with extensive manual curation, revealed only 1 and 2 receptors in the NFM transcriptome, respectively. The major limitation of the study is that, the transcriptome data of NFM is generated from a pooled sample comprising all life-stages and sexes, and thus their transcripts are likely to be highly diluted in the pooled sample. It is also important to note that the number of candidate chemosensory receptors identified in NFM is similar to that found in a recent transcriptome analysis of chemosensory gene expression in different organs in PRM (Bhowmick et al., 2020), but considerably smaller than that described in other free-living mite and tick species, as well as in insects (Figure 3). Relying on only transcriptomic data, however, may not be suitable to assess the presence or absence of IRs and GRs. Some of these receptors including GPCRs (G-protein coupled receptor) are difficult to identify in the transcriptome assembly because they are present in specialized sensory tissues (e.g. Haller's organ in ticks and foretarsal sensory organ in mites) and lower level of gene expression profiles (Bhowmick et al., 2020; Pietrantonio et al., 2018; Vizuela et al., 2020b).

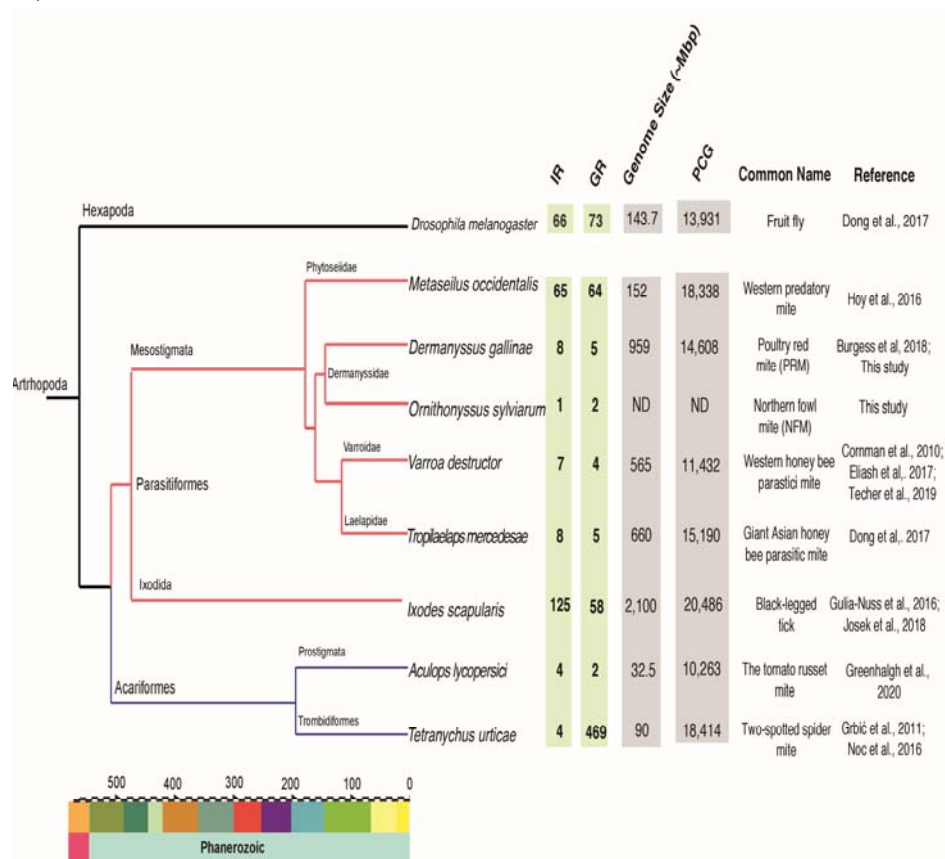


Figure 3: The number of candidate chemosensory receptor genes, genome and transcriptome statistics of poultry mites and other arthropods species. To compare the differences of IR and GR gene numbers among various mite and tick species, we compared two species of poultry mite chemoreceptor genes to those of three other Mesostigmata species and those of three additional Acari and an insect. Species are depicted with evolutionary relationships and divergence times based on Lozano-Fernandez et al. (2020) and Howard et al. (2020), with the timescale axis representing millions of years before the present. PCG: protein coding genes; ND: not determined; Mbp: million base pair.

To assess if the specialized lifestyle (extreme host specificity) of NFM has had a notable impact on the presence and abundance of IRs and GRs, we included the complete genome of the PRM, which has a similar lifestyle, in further analysis, and compared this dataset with the well-annotated genome of *I. scapularis*, which exhibits a different lifestyle (a broad range of hosts, including humans) (Figure 3). Nine GRs and forty-one iGluRs/IRs genes were annotated in the PRM genome, which have increased the completely new chemosensory gene regions in the NCBI database for this species. Accurate gene annotation from newly sequenced genomes is an important step for downstream functional and evolutionary analyses. Here, we used the BITACORA package to identify and predict all chemoreceptor family members, and then validated the chemosensory genes via ORCAE. The ORCAE-lead platform provides full annotation information about the gene models, such as gene structures, location of genes in the genome, visualization of gene expression atlas, protein homologs and domains, and definition of gene function. Each gene ID was assigned a unique loci identifier with the following format: DEGALXgY, where DEGAL denotes the species name (*D. gallinae*), X indicates the scaffold ID and Y defines the specific location within the scaffold. The predicted gene models were generated using both a computational stand-alone pipeline and manual curation, along with global gene expression profile or EST alignment data from ORCAE, which means that almost all olfactory gene loci were identified at correct locations (Figure 4). As such, annotation of the *D. gallinae* draft genome can be further improved via a process of community-led manual curation.



Figure 4: An overview of gene pages in the ORCAE resource. These pages display an extensive graphical representation of gene structure, function, coding sequence (CDS), orthologous information from other public databases, protein domains for a given gene, as well as other evidential information such as expressed sequence tag (EST) and gene expression profiles. The gene-specific page for the gustatory receptor (GR), DEGal6532g00230 is displayed. DEGal means the species name (*Dermanyssus gallinae*), 6532 denotes the scaffold ID and 00230 defines the specific location within the scaffold. Expression of this GR gene was supported by EST data.

Sequence and phylogenetic analyses of candidate IRs/iGluRs and GRs

Out of nine candidate genes encoding IR/iGluRs genes in *O. sylviarum*, six unigenes contained the specific domain signature of the ligand-gated ion channel (LCD-PF00060), which is characteristic of most IR/iGluRs subfamilies (Croset et al., 2010). Only two candidate unigenes had an amino terminal domain (ATD-PF01094), which is characteristic of traditional iGluRs (AMPA, NMDA and kainate receptors) (Croset et al., 2010). In *D. gallinae*, a total of 41 IRs/iGluRs unigenes were predicted in the genome assembly, of which 29 unigenes contained LCD domain. To obtain additional evidence supporting the homology of the conserved family of iGluR, we carried out *ab initio* 3D structure predictions of Osyl_iGluR7 (AMPA-type).

375 Osyl_iGluR7 has a unique modular architecture with four distinct domains: (1) an
ATD, which is structurally and functionally the most divergent site among the iGluR
subunits, is involved in receptor assembly, trafficking and modulation; (2) a
ligand-binding domain (LBD), which is involved in agonist and/or antagonist
recognition to activate ion channel; (3) a transmembrane domain (TMD), which
forms the membrane-spanning ion channel; and (4) an intracellular C-terminal
domain (CTD) involved in regulating synaptic efficiency, receptor mobility, and
380 trafficking (Figure 5). In order to classify the IR/iGluR family *O. sylviarum* and *D.*
gallinae, we inferred a maximum-likelihood phylogenetic tree based on
domain-specific region analysis (PF00060), which identified four major clades. Of the
seven candidate IR/iGluRs unigenes, with a conserved domain, four (Osyl_iGluR1,
Osyl_iGluR2, Osyl_iGluR3, Osyl_iGluR4) were members of the kainate receptor
subfamily, one (Osyl_iGluR7) was phylogenetically clustered with AMPA receptors,
385 and one (Osyl_iGluR8) clustered in an NMDA receptor subfamily (Figure 6). A similar
number of candidate iGluR unigenes were identified in the genome of *D. gallinae*: 10
kainate-type, 5 AMPA-type and 2 NMDA-type receptors (Figure 6). Our
phylogenetic analysis also revealed more divergent Kainate-type orthologs among
390 three classical iGluRs. This expansion is possibly derived from gene duplications.

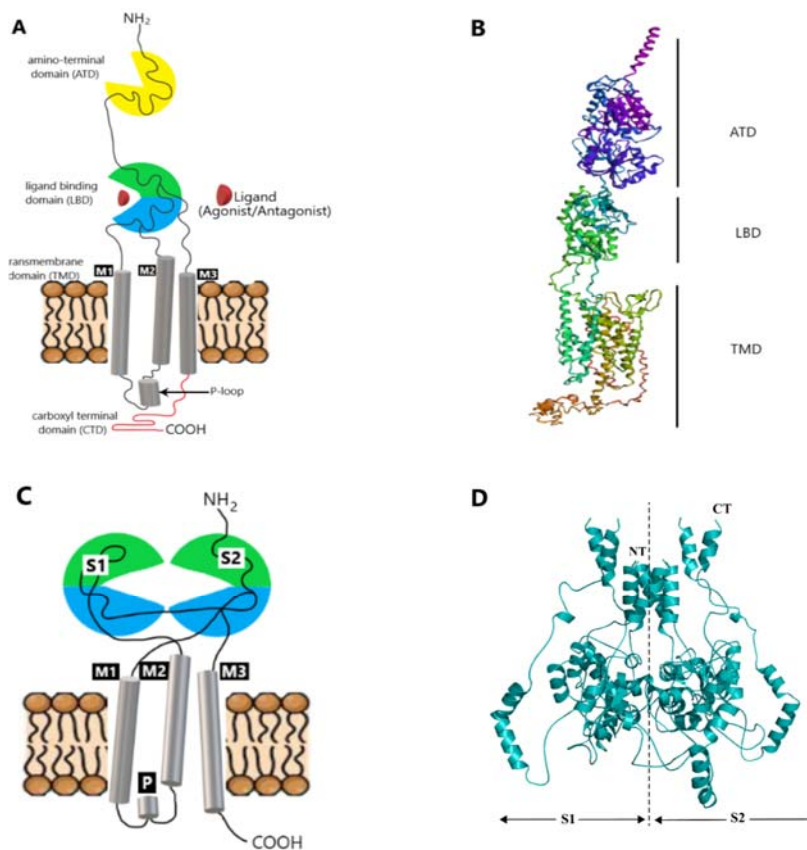


Figure 5: Overall topology of the ionotropic glutamate receptor (iGluR) and

ionotropic receptor (IR). (A) A schematic representation of the iGluR protein is shown
395 in cartoon form. The two halves of LBD (S1 and S2) are present; M1, P loop, M2, and
M3 make up the transmembrane (7_{tm}) region of each subunit. (B) The homology
model for the 3D structure of Osyl_IR7 (AMPA-type iGluR) was created via I-Tasser.
Predicted structure of the putative AMPA-type receptor demonstrates distinct domains,
including the amino terminal domain (ATD), the ligand-binding domain (LBD) and the
400 transmembrane domain (TMD). The intracellular carboxyl-terminal domain (CTD) is
not shown. (C) A general schematic representation of IR protein is demonstrated in
cartoon form. ATD is absent in most IRs. The two halves of LBD (S1 and S2) are
present. (D) The homology model for the 3D structure of Osyl_IR9 (IR25a-like) is
shown. The bilobal domain architecture is marked by the dashed line, with the S1
405 domain in the left and the S2 domain in the right side. NT means N-terminal region,
and CT indicates C-terminal region.

Like iGluRs, IRs are characterized by a presence of short extracellular
N-terminal region, a bipartite LBD (S1 and S2), LCD and a short C-terminal region.
However, our *ab initio* protein structure prediction from the amino acid sequence
410 (Osyl_IR9; IR25a-like) did not identify LCD because the homologous proteins that
have not been solved experimentally (Figure 5). In spite of having a similar domain
architecture to iGluRs, amino acid sequence similarities between IRs and iGluRs is
quite low (less than 34%), particularly within the LBD. Most IRs that lack an ATD
require a co-receptor to function (Rytz et al., 2013; van Giesen et al., 2017). Apart
415 from the iGluRs, the conserved co-receptor IR25a was identified in both *O. sylviarum*
and *D. gallinae*, which is the oldest member of the IRs (Croset et al., 2010) (Figure 6).
Interestingly, IR25a from both species shows high sequence similarity with IR25a
identified in distantly related species, e.g. the vinegar fly *D. melanogaster*. Orthologs
of the other conserved IR co-receptors for olfactory function, IR8a and IR76b (Croset
420 et al., 2010; Vizueta et al., 2018), that are present in the genomes of most insects and
non-insect arthropods, were not identified in *O. sylviarum*, the phylogenetic analyses
confirmed the presence of two putative IR93a members in *D. gallinae*, as well as 8
complete divergent IRs (Figure 6). While the transcriptome of *O. sylviarum* is highly
complete according to the BUSCO searches, the absence of co-receptors, other than
425 the IR25a ortholog, and divergent IRs in this species may be caused by the
fragmentation of sequences in the assembly, or by the lack of expression of these gene
family members in our transcriptome. The presence of a few orthologous IR
co-receptors and divergent IRs in other mites (Figure 3), along with our assessed
species, suggests either that the last common ancestor of arthropods had very few IRs,
430 or that IRs may have been lost due to different adaptations. In support of this,
Parasitiformes mites that are exposed to different types of environments and multiple
host range, such as in the case of the predatory phytoseiid mite and the black-legged
tick, have a higher number of IRs than the more strictly host-specific poultry mites
(Figure 3).

435

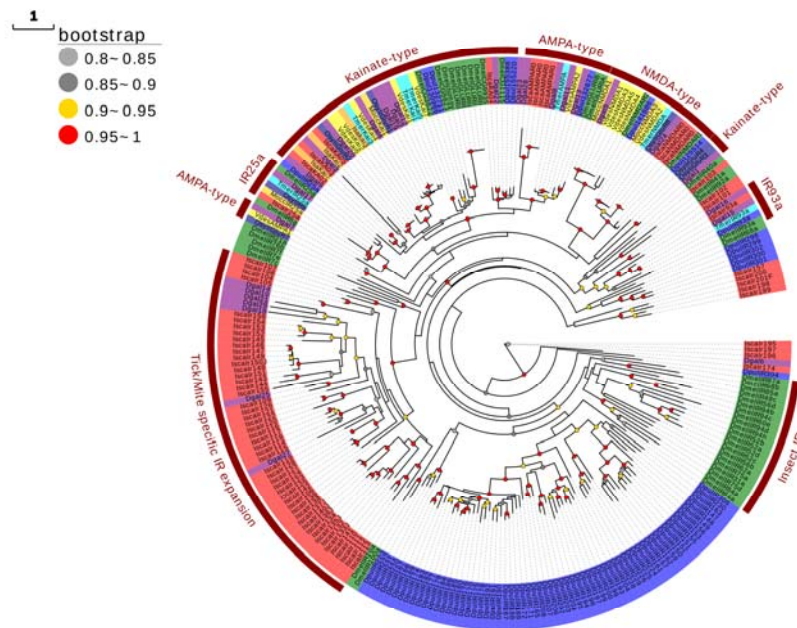
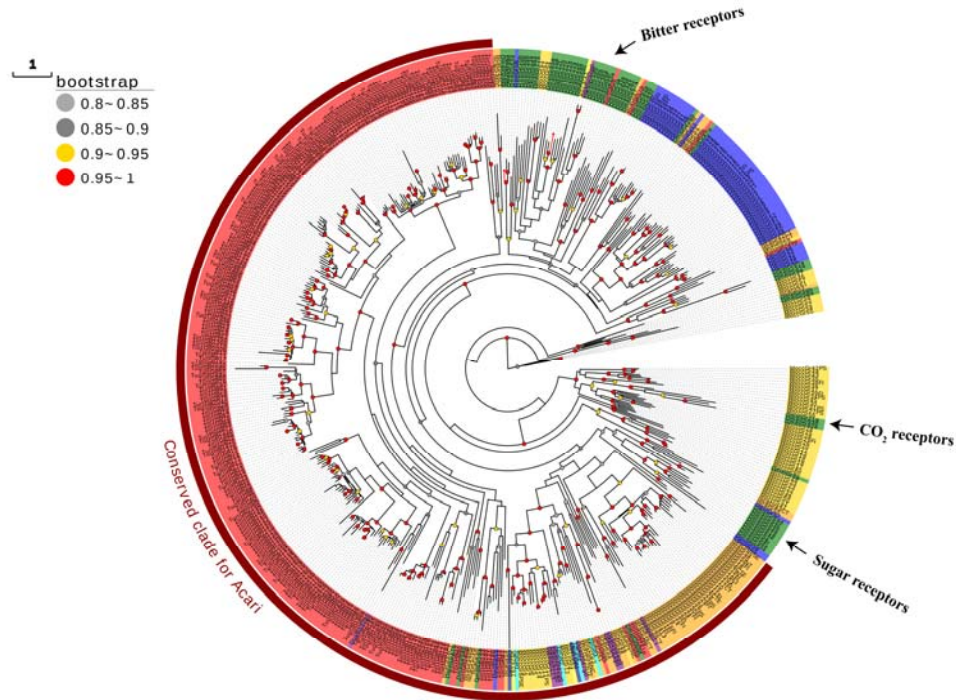


Figure 6: An evolutionary tree of the iGluR/IR gene family is based on LCD domain. This analysis involved 265 protein sequences from *Ornithonyssus sylviarum* (dark blue: Osyl), *Dermanyssus gallinae* (purple: Dgal), *Varroa destructor* (yellow: Vd),
440 *Metaseiulus occidentalis* (gold: Mocc), *Ixodes scapularis* (red: Isca), *Tropilaelaps mercedesae* (cyan: Tmer), *Drosophila melanogaster* (green: Dmel) and *Daphnia pulex* (blue: Dpul). The sequences that contain the relevant conserved domain (PF00060) were retrieved and used to construct phylogenetic analysis. The scale bar indicates amino acid changes per site.

445 Two candidate GRs, with complete ORFs containing the seven transmembrane domains (PF08395), were identified in the *O. sylviarum* transcriptome, whereas a total of nine GR genes, presenting either the 7m_7 (PF08395) or trehalose (PF06151) domains (Zhang et al., 2011), were found in the PRM genome. These candidate genes were used in phylogenetic analysis, together with representatives of the three most
450 reported GR gene family, i.e., those involved in the detection of CO₂, bitters and sugars (Zhang et al., 2011). While we were unable to identify any orthologues of these GR within these lineages in poultry mites, this may not be surprising as these genes are barely identifiable in nono-holometabolous insect lineages, including the dampwood termites *Zootermopsis nevadensis* and the damselflies *Calopteryx splendens* (Terrapon
455 et al., 2014; Ioannidis et al., 2017) (Figure 7). Most of the mite GRs clustered with GRs of ticks, suggesting that the GR repertory in these organisms originated from gene expansions specific to Acari. This suggests that Acari might use phylogenetically divergent GRs for taste detection. Chelicerates (e.g., spiders, ticks, scorpions, and mites) have evolved a lineage-specific set of GR expansions not shared with other

460 arthropods, as expected from their fast evolutionary rates and rapid sequence evolution (Vizueta et al., 2018). Interestingly, among the nine GRs in the *D. gallinae*, one GR (DEGAL5202g00060.0) was grouped with bitter receptor from *D. melanogaster* with strong bootstrap support (\square 0.95). Future functional studies are needed in order to shed light on the specific functions and evolution of chemoreceptors in Acari.



465

Figure 7: A phylogenetic tree of the GR proteins across arthropods is shown. This analysis involved 518 GR protein sequences from *Ornithonyssus sylviarum* (dark blue: Osyl), *Dermanyssus gallinae* (purple: Dgal), *Varroa destructor* (yellow: Vd), *Tetranychus urticae* (red: Turt), *Metaseiulus occidentalis* (gold: Mocc), *Ixodes scapularis* (orange: Isca), *Tropilaelaps mercedesae* (cyan: Tp), *Drosophila melanogaster* (green: Dmel) and *Daphnia pulex* (blue: Dpul). The scale bar indicates amino acid changes per site.

470

Fewer chemosensory genes are associated with the life history of poultry mites

Both *O. sylviarum* and *D. gallinae* have significantly lower numbers of predicted IR and GR genes in comparison with *I. scapularis* (Figure 3). This is likely to be explained by their simple chemosensory system of poultry mites either due to their "lifestyle as obligate parasites", or the "result of a lack of duplication". The former describes the evolutionary force that possibly shaped the genomic changes, whereas the latter refers to a molecular mechanism through which the changes may occur. This is in line with that described for the the human body louse, *Pediculus humanus humanus*, the bed bug, *Cimex lectularius* and the tomato russet mite, *Aculops lycopersici* (TRM; the world's smallest plant-eating mites), which all have a

475

480

low number of genes associated with chemoreception as well as detoxification (Kirkness et al., 2010; Benoit et al., 2016; Greenhalgh et al., 2020). Alternatively, the
485 small chemosensory gene set may be a result of a lack of duplication (fewer expansions of presumably small ancestral gene families repertoires), which previously has been demonstrated in *T. mercedesae*, which has undergone the fewest gene family expansion or contraction events since its divergence from the last common ancestor of arthropods (Dong et al., 2017). A lower number of genes associated with
490 environmental sensing have also been shown to be present in other obligate parasites, compared to closely related free-living relatives in flies, ticks and beetles (Kirkness et al., 2010). As a whole, comparative genomics and transcriptomics analysis offer unique information and tools to use in advancing the understanding of host-parasite co-evolution and signatures of adaptive evolution. Further functional studies combined
495 with new evidence based on greater coverage of phylogenetic analyses can greatly contribute to understand the chemosensory gene evolution.

Conclusions

In conclusion, the *O. sylviarum* transcriptomic data and annotation results can provide a useful resource for future studies in this non-model mite species, which is
500 an economically important pest. This study improves our understanding of its molecular biology that may result in an improved integrated pest management program, providing a tool for further identification and characterization of genes of interest and development of eco-friendly acaricides or mite repellents. Additionally, a comprehensive analysis of variant and conserved receptors in two species of poultry
505 mites revealed many features that will help in elucidating the molecular basis of the function and regulation of the chemosensory system in mites. The decreased number of chemosensory receptors predicted in bird mites may reflect the unique adaptations of these species to their specific behavioral and reproductive lifestyle.

510 Supplementary material

The Supplementary Material for this article is provided.

Ethics statement

Not applicable.

515

Author contributions

BB and QH conceived and designed the study. BB and HC wrote the first draft and performed the experiments. JL, JV, RI, and QH commented on study design, methodology, and substantially revised the manuscript. All authors approved the final
520 version.

Acknowledgments

We would like to thank Yu Tang, Dr. Jianguo Zhao, and Dr. Chenghong Liao for their assistance. This work was supported by the National Key Research and Development
525 Program of China (2017YFD0501702-9) and Hainan Provincial Key R & D Program

(ZDYF2019073).

Conflicts of Interest

None.

530

References

1. Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69, 44–60. doi: 10.1016/j.neuron.2010.11.042.
- 535 2. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402. doi: 10.1093/nar/25.17.3389.
- 540 3. Benoit, J. B., Adelman, Z. N., Reinhardt, K., Dolan, A., Poelchau, M., Jennings, E. C., et al. (2016). Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nat Commun.* 7:10165. doi: 10.1038/ncomms10165.
- 545 4. Bernsel, A., Viklund, H., Hennerdal, A., Elofsson, A. (2009). TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res.* 37, W465-468. doi: 10.1093/nar/gkp363.
5. Benton, R., Vannice, K. S., Gomez-Diaz, C., Vosshall, L. B. (2009). Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* 136, 149–162. doi: 10.1016/j.cell.2008.12.001
- 550 6. Bhowmick, B., Zhao, J., Øines, Ø., Bi, T., Liao, C., Zhang, L., et al. (2019). Molecular characterization and genetic diversity of *Ornithonyssus sylviarum* in chickens (*Gallus gallus*) from Hainan Island, China. *Parasit. Vectors.* 12:553. doi: 10.1186/s13071-019-3809-9.
- 555 7. Bhowmick, B., Tang, Y., Lin, F., Øines, Ø., Zhao, J., Liao, C., et al. (2020). Comparative morphological and transcriptomic analyses reveal chemosensory genes in the poultry red mite, *Dermanyssus gallinae*. *Sci. Rep.* 10:17923. doi: 10.1038/s41598-020-74998-7.
- 560 8. Burgess, S. T. G., Marr, E. J., Bartley, K., Nunn, F. G., Down, R. E., Weaver, R. J., et al. (2019). A genomic analysis and transcriptomic atlas of gene expression in *Psoroptes ovis* reveals feeding- and stage-specific patterns of allergen expression. *BMC Genomics* 20:756. doi: 10.1186/s12864-019-6082-6.
- 565 9. Burgess, S. T. G., Bartley, K., Nunn, F., Wright, H.W., Hughes, M., Gemmell, M., et al. (2018). Draft Genome Assembly of the Poultry Red Mite, *Dermanyssus gallinae*. *Microbiol Resour Announc.* 7:e01221–01218. doi: 10.1128/MRA.01221-18.
10. Clyne, P. J., Warr, C. G., Carlson, J. R. (2000). Candidate taste receptors in *Drosophila*. *Science* 287, 1830–1834.
11. Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L., Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina

- 570 FASTQ variants. *Nucleic Acids Res.* 38, 1767-17671. doi: 10.1093/nar/gkp1137.
12. Cornman, S. R., Schatz, M. C., Johnston, S. J., Chen, Y. P., Pettis, J., Hunt, G., et al. (2010). Genomic survey of the ectoparasitic mite *Varroa destructor*, a major pest of the honey bee *Apis mellifera*. *BMC Genomics* 11:602. doi: 10.1186/1471-2164-11-602.
- 575 13. Chyb, S., Dahanukar, A., Wickens, A., Carlson, J. R. (2003). *Drosophila* Gr5a encodes a taste receptor tuned to trehalose. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14526–14530. doi: 10.1073/pnas.2135339100
14. Croset, V., Rytz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H.,
580 (2010). Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* 6:e1001064. doi: 10.1371/journal.pgen.1001064.
15. Dong, X., Armstrong, S. D., Xia, D., Makepeace, B. L., Darby, A. C., Kadowaki, T. (2017). Draft genome of the honey bee ectoparasitic mite, *Tropilaelaps mercedesae*, is shaped by the parasitic life history. *Gigascience* 6:1-17. doi: 10.1093/gigascience/gix008.
- 585 16. Eliash, N., Singh, N. K., Thangarajan, S., Sela, N., Leshkowitz, D., Kamer, Y., et al. (2017). Chemosensing of honeybee parasite, *Varroa destructor*: Transcriptomic analysis. *Sci. Rep.* 7:13091. doi: 10.1038/s41598-017-13167-9.
17. Enjin, A., Zaharieva, E. E., Frank, D. D., Mansourian, S., Suh, G. S. B., Gallio, M., et al. (2016). Humidity sensing in *Drosophila*. *Curr. Biol.* 26, 1352–1358. doi: 10.1016/j.cub.2016.03.049.
- 590 18. Eyun, S. I., Soh, H. Y., Posavi, M., Munro, J. B., Hughes, D. S. T., Murali, S. C., et al. (2017). Evolutionary History of Chemosensory-Related Gene Families across the Arthropoda. *Mol. Biol. Evol.* 34, 1838-1862. doi: 10.1093/molbev/msx147.
- 595 19. Faraone, N., Light, M., Scott, C., MacPherson, S., Hillier, N. K. (2020). Chemosensory and Behavioural Responses of *Ixodes scapularis* to Natural Products: Role of Chemosensory Organs in Volatile Detection. *Insects* 11:502. doi: 10.3390/insects11080502.
- 600 20. Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2015). The Pfam protein families database: towards a more sustainable future. *Nucl. Acids. Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344.
21. Gay, M., Lempereur, L., Francis, F., Caparros Megido, R. (2020). Control of *Dermanyssus gallinae* (De Geer 1778) and other mites with volatile organic
605 compounds, a review. *Parasitology* 147, 731–739. doi: 10.1017/S0031182020000530.
22. Greenhalgh, R., Dermauw, W., Glas, J. J., Rombauts, S., Wybouw, N., Thomas, J., et al. (2020). Genome streamlining in a minute herbivore that manipulates its host plant. *eLife* 9:e56689. doi: 10.7554/eLife.56689.
- 610 23. Grbić, M., Van Leeuwen, T., Clark, R. M., Rombauts, S., Rouzé, P., Grbić, V., et al. (2011). The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* 479, 487-492. doi: 10.1038/nature10640.
24. Guo, Z., Qin, J., Zhou, X., Zhang, Y. (2018). Insect Transcription Factors: A

- Landscape of Their Structures and Biological Functions in *Drosophila* and beyond. *Int. J. Mol. Sci.* 19:3691. doi: 10.3390/ijms19113691.
- 615 25. Gulia-Nuss, M., Nuss, A. B., Meyer, J. M., Sonenshine, D. E., Roe, R. M., Waterhouse, R. M., et al. (2016). Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nat. Commun.* 7:10507. doi: 10.1038/ncomms10507.
26. He, Z., Zhang, H., Gao, S., Lercher, M. J., Chen, W. H., Hu, S. (2016). Evolview
620 v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* 44, W236-W41. doi: 10.1093/nar/gkw370.
27. Howard, R. J., Puttick, M. N., Edgecombe, G. D., Lozano-Fernandez, J. Arachnid monophyly: Morphological, palaeontological and molecular support for a single
625 terrestrialization within Chelicerata. *Arthropod Struct Dev.* 2020 Nov;59:100997. doi: 10.1016/j.asd.2020.100997.
28. Hoy, M. A. (2009). The predatory mite *Metaseiulus occidentalis*: mitey small and mitey large genomes. *BioEssays* 31, 581–590. doi: 10.1002/bies.200800175.
29. Hoy, M. A., Waterhouse, R. M., Wu, K., Estep, A. S., Ioannidis, P., Palmer, W. J.,
630 et al. (2016). Genome Sequencing of the Phytoseiid Predatory Mite *Metaseiulus occidentalis* Reveals Completely Atomized Hox Genes and Superdynamic Intron Evolution. *Genome Biol. Evol.* 8, 1762-1775. doi: 10.1093/gbe/evw048.
30. Huang, Y., Li, H., Wang, C., Xu, X., Yu, H., Meng, J., et al. (2020). De novo assembly and discovery of genes related to blood digestion in the transcriptome of *Dermanyssus gallinae* (Acari: Dermanyssidae). *Vet. Parasitol.* 286:109246.
635 doi: 10.1016/j.vetpar.2020.109246.
31. Ioannidis, P., Simao, F.A., Waterhouse, R.M., Manni, M., Seppey, M., Robertson, H.M., et al. (2017). Genomic features of the damselfly *Calopteryx splendens* representing a sister clade to most insect orders. *Genome Biol Evol.* 9, 415–430. doi: 10.1093/gbe/evx006.
- 640 32. Jones, W. D., Cayirlioglu, P., Kadow, I. G., Vosshall, L. B. (2007). Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature* 445, 86–90. doi: 10.1038/nature05466.
33. Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome biology* 5:226. <https://doi.org/10.1186/gb-2004-5-6-226>
- 645 34. Joseph, R. M., Carlson, J. R. (2015). *Drosophila* Chemoreceptors: A Molecular Interface Between the Chemical World and the Brain. *Trends Genet.* 31, 683-695. doi: 10.1016/j.tig.2015.09.005.
35. Josek, T., Walden, K. K. O., Allan, B. F., Alleyne, M., Robertson, H. M. (2018). A foreleg transcriptome for *Ixodes scapularis* ticks: Candidates for
650 chemoreceptors and binding proteins that might be expressed in the sensory Haller's organ. *Ticks Tick Borne Dis.* 9, 1317-1327. doi: 10.1016/j.ttbdis.2018.05.013.
36. Josek, T., Sperrazza, J., Alleyne, M., Syed, Z. (2021). Neurophysiological and behavioral responses of blacklegged ticks to host odors. *J. Insect Physiol.* 128:104175. doi: 10.1016/j.jinsphys.2020.104175.
- 655 37. Katoh, K., Rozewicki, J., Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief

- Bioinform. 20, 1160-1166. doi: 10.1093/bib/bbx108.
- 660 38. Kirkness, E. F, Haas, B. J., Sun, W., Braig, H. R., Perotti, M. A., Clark, J. M., et al. (2010). Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc. Natl. Acad. Sci. U. S. A.* 107, 12168-12173. doi: 10.1073/pnas.1003379107.
39. Lefort, V., Longueville, J. E., Gascuel, O. (2017). SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* 34, 2422-2424. doi: 10.1093/molbev/msx149.
- 665 40. Leite, D. J., Baudouin-Gonzalez, L., Iwasaki-Yokozawa, S., Lozano-Fernandez, J., Turetzek, N., Akiyama-Oda, Y., et al. (2018). Homeobox Gene Duplication and Divergence in Arachnids. *Mol. Biol. Evol.* 35, 2240-2253. doi: 10.1093/molbev/msy125.
41. Light, M., Shutler, D., Cutler, G. C., Hillier, N. K. (2020). *Varroa destructor* mite electrophysiological responses to honey bee (*Apis mellifera*) colony volatiles. *Exp. Appl. Acarol.* 81, 495-514. doi: 10.1007/s10493-020-00519-w.
- 670 42. Lozano-Fernandez J, Tanner AR, Puttick MN, Vinther J, Edgecombe GD and Pisani D (2020) A Cambrian–Ordovician Terrestrialization of Arachnids. *Front. Genet.* 11:182. doi: 10.3389/fgene.2020.00182.
- 675 43. Murillo, A. C., Abdoli, A., Blatchford, R. A., Keogh, E. J., Gerry, A. C. (2020). Parasitic mites alter chicken behaviour and negatively impact animal welfare. *Sci. Rep.* 10:8236. doi: 10.1038/s41598-020-65021-0.
44. Ngoc, P. C., Greenhalgh, R., Dermauw, W., Rombauts, S., Bajda, S., Zhurov, V., et al. (2016). Complex Evolutionary Dynamics of Massively Expanded Chemosensory Receptor Families in an Extreme Generalist Chelicerate Herbivore. *Genome Biol. Evol.* 8, 3323–3339. doi: 10.1093/gbe/evw249.
- 680 45. Ni, L., Klein, M., Svec, K. V., Budelli, G., Chang, E. C., Ferrer, A. J., et al. (2016). The ionotropic receptors IR21a and IR25a mediate cool sensing in *Drosophila*. *eLife* 5:e13254. doi: 10.7554/eLife.13254
- 685 46. Pietrantonio, P. V., Xiong, C., Nachman, R. J., Shen, Y. (2018). G protein-coupled receptors in arthropod vectors: omics and pharmacological approaches to elucidate ligand-receptor interactions and novel organismal functions. *Curr. Opin. Insect Sci.* 29, 12-20. doi: 10.1016/j.cois.2018.05.016.
47. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic Acids Res.* 2018 Jul 2;46(W1):W200-W204. doi: 10.1093/nar/gky448.
- 690 48. Roy, L., Dowling, A. P. G., Chauve, C. M., Lesna, I., Sabelis, M. W., Buronfosse, T. (2009). Molecular phylogenetic assessment of host range in five *Dermanyssus* species. *Exp. Appl. Acarol.* 48:115. doi: 10.1007/s10493-008-9231-1.
- 695 49. Roy A., Kucukural, A., Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5: 725-738.
- 700 50. Rytz, R., Croset, V., Benton, R. Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. (2013). *Insect Biochem Mol Biol.* 43, 888-897. doi: 10.1016/j.ibmb.2013.02.007.

51. Schicht, S., Qi, W., Poveda, L., Strube, C., 2014. Whole transcriptome analysis of the poultry red mite *Dermanyssus gallinae* (De Geer, 1778). *Parasitol* 141, 336–346.
- 705 52. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. (2015). *Bioinformatics* 31, 3210-3212. doi: 10.1093/bioinformatics/btv351.
53. Sonenshine, D. E. Pheromones and other semiochemicals of ticks and their use in tick control. *Parasitology* 2004, 129, S405–S425.
- 710 54. Sobolevsky, A., Rosconi, M. & Gouaux, E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* 462, 745–756 (2009).
55. Sparagano, O. A. E., Ho, J., 2020. Parasitic Mite Fauna in Asian Poultry Farming Systems. *Front. Vet. Sci.* 7, 400.
- 715 56. Sterck, L., Billiau, K., Abeel, T., Rouzé, P., Van de Peer, Y.. (2012). ORCAE: online resource for community annotation of eukaryotes. *Nat. Methods* 9:1041. doi: 10.1038/nmeth.2242. PMID: 23132114.
57. Techer, M. A., Rane, R. V., Grau, M. L., Roberts, J. M. K., Sullivan, S. T., Liachko, I., et al. (2019). Divergent evolutionary trajectories following speciation in two ectoparasitic honey bee mites. *Commun Biol.* 2:357. doi: 10.1038/s42003-019-0606-0.
- 720 58. Tomley, F. M., Sparagano, O. (2018). Spotlight on avian pathology: red mite, a serious emergent problem in layer hens. *Avian Pathol.* 47, 533-535. doi: 10.1080/03079457.2018.1490493.
- 725 59. Terrapon, N., Li, C., Robertson, H.M., Ji, L., Meng, X., Booth, W., Chen, Z., Childers, C.P., Glastad, K.M., Gokhale, K., Gowin, J., Gronenberg, W., Hermansen, R.A., Hu, H., et al., 2014. Molecular traces of alternative social organization in a termite genome. *Nat. Commun.* 5, 3636.
- 730 60. van Giesen, L., Garrity, P. A. (2017). More than meets the IR: the expanding roles of variant Ionotropic Glutamate Receptors in sensing odor, taste, temperature and moisture. *F1000Res.* 6:1753. doi: 10.12688/f1000research.12013.1.
61. Vizueta, J., Rozas, J., Sánchez-Gracia, A. (2018). Comparative Genomics Reveals Thousands of Novel Chemosensory Genes and Massive Changes in Chemoreceptor Repertoires across Chelicerates. *Genome biology and evolution*, 10(5), 1221–1236.
- 735 62. Vizueta, J., Escuer, P., Frías-López, C., Guirao-Rico, S., Hering, L., Mayer, G, et al. (2020a). Evolutionary History of Major Chemosensory Gene Families across Panarthropoda. *Mol. Biol. Evol.* 37, 3601-3615. doi: 10.1093/molbev/msaa197.
- 740 63. Vizueta, J, Sánchez-Gracia, A, Rozas, J. (2020b). BITACORA: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Mol. Ecol. Resour.* 20, 1445-1452. doi: 10.1111/1755-0998.13202.
64. Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., Axel, R. (1999). A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* 96,
- 745

725–736. doi: 10.1016/s0092-8674(00)80582-6.

65. Yssel, A. E. J., Kao, S. M., Van de Peer, Y., Sterck, L. (2019). ORCAE-AOCC: A Centralized Portal for the Annotation of African Orphan Crop Genomes. *Genes* 10:950. doi: 10.3390/genes10120950.
- 750 66. Zhang, H. J., Anderson, A. R., Trowell, S. C., Luo, A. R., Xiang, Z. H., Xia, Q. Y. (2011). Topological and functional characterization of an insect gustatory receptor. *PLoS One*. 6:e24111. doi: 10.1371/journal.pone.0024111.