The Genome of Rice Planthopper Egg Parasitoid Wasps *Anagrus Nilaparvatae* (Hymenoptera: Mymaridae) Casts Light on the Chemo- and Mechano-Sensing in Parasitism

Ying Ma

Sun Yat-Sen University

Zixiao Guo

Sun Yat-Sen University

Liyang Wang

Sun Yat-Sen University

Bingyang Wang

Sun Yat-Sen University

Tingfa Huang

Sun Yat-Sen University

Bingjie Tang

Sun Yat-Sen University

Guren Zhang

Sun Yat-Sen University

Qiang Zhou (✉ lsszhou@mail.sysu.edu.cn)

Sun Yat-Sen University

Research Article

Keywords: *Anagrus nilaparvatae*, Phylogenetic analysis, Olfactory receptor protein, Taste receptor protein, Mechanoreceptor protein

Posted Date: March 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1394033/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Mymaridae is an ancient insect group, which is a basal lineage in the Chalcidoidea phylogeny. The species of Mymaridae have great potential for biological control. *Anagrus nilaparvatae*, a representative species of Mymaridae, is ideal for controlling rice planthopper due to its high rate of parasitism by finding hosts efficiently from paddy ridges to fields.

Results

Using both single-molecule real-time and Illumina sequencing, we sequenced and assembled the whole genome of *A. nilaparvatae*, which is the first genome sequenced in family Mymaridae. The assembly consists of 394 scaffolds accumulated to 488.8Mb. The assembly is of high continuity and completeness, indicated by the N50 value of 25,368,259bp and 98.4% mapping rate of BUSCO orthologs. 16894 protein-coding genes in the genome were annotated. The phylogenomic tree constructed for *A. nilaparvatae* and other 12 species of Hymenoptera confirmed that family Mymaridae is basal in the Chalcidoidea superfamily and sister to all the remaining Chalcidoidea species. The divergence between *A. nilaparvatae* and the other seven Chalcidoidea species was dated at ~ 126.9 Mya. Chemoreceptor and mechanoreceptor genes are important in the parasitic behavior. We identified 17 OBPs, 11 CSPs, 4 NPC2s, 88 ORs, 7 GRs, 22 IRs and 13 SNMPs in the genome of *A. nilaparvatae*, which are associated with the chemosensory function. Strikingly, there are only 1 PPK and 3 TRP genes in the genome that are relevant to the mechanosensory function.

Conclusions

The small numbers of chemo- and mechano-sensory genes in *A. nilaparvatae* indicate the species-specific host detection and oviposition behavior of *A. nilaparvatae* might be regulated by relatively simple molecular pathways.

Background

Parasitic wasps are in the spotlight for the potential of biological control of crop pests. In the long process of coevolution with host, parasitic wasps have formed unique parasitic behaviors such as oviposition recognition and post-oviposition processing [1]. In the process of parasitism, parasitic wasps recognize host-related chemical information materials through smell and taste, and recognize the protective traits of hosts through touch. Facilitated by smell, taste, touch, and other senses, parasitic wasps accurately anchor at the appropriate position on the host pests to complete the parasitic behavior [2].

The superfamily Chalcidoidea is a large group, the majority of which are parasitoid wasps. The number of parasitoid wasp species described so far is more than 500,000, accounting for about 75% of the total Hymenoptera species and 10–20% of the total insect species [3]. Mymaridae is an important family in the Chalcidoidea superfamily. Previous diagnostics using phylogenetic analyses, morphological characteristics, molecular data and fossil evidences suggest that Mymaridae is an ancient insect group, which is a basal lineage in the Chalcidoidea phylogeny and sister to all the other Chalcidoidea lineages [4–6]. Fossils of the earliest Chalcidoidea suggest that the ancestors of Chalcidoidea might have been small egg parasites [6], and Mymaridae has retained the characteristics of egg parasitism.

Rice planthoppers are the most damaging insect pests in rice producing areas in Asia. The wasps of *Anagrus nilaparvatae* Pang et Wang (Hymenoptera: Mymaridae) are parasitic in the eggs of the planthoppers and are the dominant natural enemy of rice planthoppers, with great application value in rice production [7, 8]. The parasitic wasps spend the winter in paddy ridges, parasitize the rice planthoppers in the paddy fields in the spring, and return to the ridges after the rice is harvested (Fig. 1). In the process of finding a host, *A. nilaparvatae* wasps recognize the volatiles released by rice plants that are hurt by rice planthoppers [9–11]. They continuously tap the stem of rice plants to anchor the position of host eggs, and they perceive the inner texture of the rice track using the ovipositor to distinguish the host eggs during the spawning process (Supplementary data). It is speculated that olfactory perception, taste perception and mechanical perception play important roles in the process of hosting.

The molecular signaling pathway of insect olfactory sensation consists of several classes of proteins functioning in different steps, including odorant binding proteins (OBPs), chemosensory proteins (CSPs), niemann-pick type C2 proteins (NPC2s), olfactory receptors (ORs), ionotropic receptors (IRs), sensory neuron membrane proteins (SNMPs), and odorant-degrading enzymes (ODEs) [12–14]. In contrast, the taste sensation is signaled by gustatory receptors (GRs) in insects [15–17]. Moreover, the mechanical perception in insects is signaled by pickpocket receptors (PPKs) and transient receptor potential (TRP) channels [18, 19]. To uncover the origin of Mymaridae parasitic wasps, and the genomic basis of their parasitic behavior, we sequenced and assembled the whole genome of *A. nilaparvatae*, which is the first genome reported in family Mymaridae. The genome assembly provide new insights for the Chalcidoidea phylogeny and enables reconstruction of the evolutionary history of gene families related to chemo- and mechano-sensory.

Results

Genomic assembly information

The genome of *A. nilaparvatae* was sequenced by the PacBio and the Illumina platform. A total of 24.6 Gb of SMRT long reads and a total of 42.2 Gb of Illumina raw reads were obtained, the average coverage of SMRT long reads is 55X (Table S1), and the average sequencing depth of Illumina reads is 95X

(Table S2). With K-mer = 55, the genome was estimated as 444 Mb, with a heterozygosity rate of 0.634% and a repetition rate of 45.5% (Fig. S1). The genome was finally assembled to 488.8 Mb. The assembly contains a total of 394 scaffolds, the largest scaffold is 70.7 Mb and the N50 size is 25.4 Mb (Table 1).

Table 1
Summary of the genome assembly of *A. nilaparvatae*

Statistics	<i>Anagrus nilaparvatae</i>
Number of scaffolds	394
Total length	488,841,863 bp
Longest scaffold	74,101,551 bp
Scaffold N50	25,368,259 bp
Scaffold N90	2,930,360 bp
GC content (%)	27.52
N content (%)	0
Gene number	16861
Average gene length	8075.1 bp
Exon number per gene	5.42
Average exon length	275.0 bp
Excon GC content (%)	32.53
Note: Scaffold N50 (N90) indicates the scaffold size which accumulates to 50% (90%) of the whole genome by ranking all scaffolds from large to small.	

We evaluated the integrity of the assembly by mapping the single-copy homologous genes in the eukaryotic database of BUSCO to the assembly. As a result, 98.4% of the 255 core genes are found in the assembly of *A. nilaparvatae* genome, with 89.0% completely mapped (Table S3). We also mapped the Illumina clean reads back to the assembly, 89.26% of the reads are successfully mapped. These assessments indicated that our assembly is with good completeness. The genomewide average GC content is 27.52% and the GC contents in scaffolds are between 20%~40% (Fig. S2).

Circos software was used to analyze 22 large scaffolds in the genome (the scaffold length > 3 Mb). There are many colinear genes between *utg1* and *utg14* and between *utg8* and *utg30* (Fig. 2).

Gene Annotation

RepeatModeler and RepeatMasker were used to mask all the repeated sequences except the low complexity sequences (0.01%) and simple repetition sequences (0.04%) from the assembly. In total, 52.92% of the genome were masked, with interspersed repeats occupying 52.48% and tandem repeats occupying 0.07% of the whole genome (Table S4). The transposable elements (TEs) in *A. nilaparvatae* are mostly long interspersed nuclear elements (LINEs, 5.61% of genome), DNA transposons (2.30% of genome) and long terminal repeat retrotransposons (LTRs, 3.61% of the genome), while short interspersed nuclear elements (SINEs) are few (Table S4; Fig. S3). By blasting against the homologous sequences in NR, COG, GO, KEGG and other databases, we predicted 16,861 protein coding genes in the genome of *A. nilaparvatae* (Table S5). The average length of the genes is 1404.8 bp and the average exon length is 279.6 bp (Table 1). Functions of the predicted genes are annotated with the KOG database (Fig. S4 and S5).

Phylogenetic analysis and gene family evolution

We used the maximum likelihood method to construct a phylogeny for *A. nilaparvatae* and 12 other species of Hymenopteras, with *A. mellifera* as outgroup (Fig. 3). The phylogeny diverged into two lineages, with one monophyletic clade of the eight species of Chalcidoidea and the clade of Ichneumonoidea and Cynipoidea. The three species of Ichneumonoidea also form a monophyly. Within the Chalcidoidea clade, *A. nilaparvatae* is the most basal lineage, sister to all the other species of Chalcidoidea. *C. solmsi* is second basal, sister to the clade of Trichogrammatidae and Encyrtidae (*T. brassicae*, *T. pretiosum*, *C. floridanum*), and the clade of Pteromalidae (*N. vitripennis*, *T. sarcophagae*, *P. puparum*). Calibrated by the known fossil data [20] and previous estimations [6], we dated the split of *A. mellifera* from the others (species of Terebrantia) at about 213.1 Mya (95% CI: 199.2-242.1). Chalcidoidea diverged from the clade of Ichneumonoidea and Cynipoidea at about 188.3 Mya (95% CI: 159.0-217.2), and the split between Ichneumonoidea and Cymaoidea is dated at about 117.8 Mya (95% CI: 94.7-143.7). The time of split between *A. nilaparvatae* and the other Chalcidoidea species was about 126.9 Mya (95% CI: 116.8–132.0).

Gene family evolution showed that gene family contraction overwhelmed expansion in most species except *F. arisanus* and *N. vitripennis*. Notably, *A. nilaparvatae* and *C. solmsi* have experienced striking contraction of gene families.

Chemo- and Mechano-sensory Related Genes

We identified 17 OBPs, 11 CSPs, 4 NPC2s, 88 ORs, 7 GRs, 23 IRs, 13 SNMPs, 1 PPK, and 3 TRPS in the genome of *A. nilaparvatae*. The detailed information of the numbers in the other 12 Hymenopteras can be found in Table 2.

Table 2
The number of candidate chemo- and mechano-sensory related genes of *Anagrus nilaparvatae* and other hymenopteran species

statistics	Mymaridae	Agaonidae	Pteromalidae			Encyrtidae	Trichogrammatidae		Braconidae			Cynipidae	Apidae
	<i>Ani</i> ²	<i>Cso</i> ²	<i>Ppup</i> ¹	<i>Tsar</i> ²	<i>Nvit</i> ¹	<i>Cflo</i> ²	<i>Tpre</i> ²	<i>Tbra</i> ³	<i>Dal</i> ²	<i>Far</i> ²	<i>Mdem</i> ²	<i>Btre</i> ¹	<i>Amel</i>
Genome size (Mb)	466.2	280.4	338.1	236.4	297.3	53.9	187.64	235.4	384.4	153.6	241.2	1538.7	225.2
	16,861	11,412	17,656	16,084	24,388	12,932	13,395	16,905	13,480	11,691	12,892	14,488	9,935
Genes (n)	17	8(7)	57	48	78(90)	27	33	4	16(15)	12	13	29	19(21)
OBP	11	8	9	7	10(9)	10	9	4	9(9)	10	9	17	9(6)
CSP	4	2	6	5	7	2	8	1	4	3	8	3	3
NPC2	88	46	145	156	333(301)	67	118	15	216(201)	218	147(218)	89	231(1)
OR	7	3	12	19	51(58)	20	14	4	13(40)	24	11(85)	26	20(10)
GR	23	29	35	36	90(19)	30	44	16	40(56)	51	35	31	34(18)
IR	13	11	12	9	25	22	29	14	18	17	18	20	24(8)
SNMP	1	4	9	10	17	9	10	6	8	14	5	8	9(8)
PPK	9	11	12	10	45(12)	14	36	9	16	23	23	36	45(13)
TRP													

Note: The assembly statistics of genomes for other hymenopteran species were obtained from NCBI (<https://www.ncbi.nlm.nih.gov/>). Gene number refers to number of candidate genes screened by our blast (e-value 10^{-10}), and the numbers in brackets were reported in other references [21–31]. Superscripts 1,2 or 3 represent assembly levels of Chromosome, Scaffold and Contig respectively.

OBPs, CSPs and NPC2s are three kinds of soluble chemoreceptor proteins, which have the function of recognizing binding and transporting chemical substances such as odor molecular pheromones. They were found that not only involved in olfactory sensation, but also play roles in developmental reproduction and anti-stress functions [32, 33]. Both CSPs and OBPs have conserved cysteine (Cys) domains and their secondary structures is multiple α -helices [34–37]. Among the investigated species, the number of OBP in the three species of Pteromalidae was the largest (48 ~ 78), followed by the 33 of *T. pretiosum*. The other species have less than 30 OBPs. According to a phylogenetic analysis, the OBP genes of the 13 Hymenoptera species are clustered into seven clades, and the 17 OBP genes of *A. nilaparvatae* occurred in four of the seven clades. Particularly, seven of the 17 OBPs of *A. nilaparvatae* are sequentially located on scaffold utg34, suggestive of an origination from tandem duplication (Fig. 4). The number of CSP genes of *A. nilaparvatae* (11) is only less than that of *B. treatae* (17), and the other species have less than 10 CSPs. The CSPs are phylogenetically clustered into six clades, with *A. nilaparvatae* CSPs occurring in four clades (Fig. S6). In Subgroup2, five CSPs of *A. nilaparvatae* are phylogenetically distant to other members in the same clade, indicating fast evolutions in *A. nilaparvatae*.

NPC2s were found to be functionally similar to OBPs in recent years [38–40]. Aligned with OBPs and CSPs, the NPC2 genes also contains the conserved cysteine domains, but the insect NPC2 proteins are mainly based on β -sheet in secondary structures and form a larger endo-binding cavity [38]. All the 13 species have less than 10 NPC2 candidate genes and four are found in the genome of *A. nilaparvatae*. The NPC2s are clustered into three clades (Fig. S7).

The three classes of soluble chemoreceptor proteins mentioned above deliver chemical pheromones or environmental odors to the chemoreceptors of sensory neurons. Subsequently, three transmembrane transporters, ORs, GRs, and IRs, are responsible for the recognition and discrimination of these signals. Both ORs and GRs belong to the G-protein-coupled receptors (GPCRs), containing seven transmembrane domains [15, 41]. GRs were initially found to mostly express in mouth and other taste organs, while ORs mainly express in antenna [15]. However, GRs were later found to express in olfactory structures, indicating they may function as olfactory receptors [16]. IRs of insects belongs to the ionotropic glutamate receptor (IGluR) family, which is a class of conservative ligand-gated ion channels. The structures of IR proteins include extracellular N-terminal (N) and ligand-binding domain (LBD, composed of S1 and S2), three transmembrane domains, ion channel pores and intracellular C-terminal [42]. In recent years, IRs have been found to be involved in taste, temperature and humidity besides its olfactory function [43–45]. SNMPs are a specific class of double transmembrane proteins on olfactory neurons and dendrite membranes, which assist ORs in the process of insect sex pheromone recognition [13, 46, 47]. SNMP is a homologous protein of the Mammalian CD36 gene family [27, 48], the representative SNMP has two transmembrane domains, c-terminal and N-terminal, and an extracellular ring including several N-glycosylated groups and disulfide bonds [49, 50].

There were a large number of OR candidate genes in each of the 13 species, and the species of Pteromalidae and Braconidae, as well as *Apis mellifer* have over 200 ORs (Fig. S8). Particularly, we identified 88 ORs in *A. nilaparvatae*, which are likely clustered into six clades. For GRs, we found all the 13 species have less than 30 GRs, except *N. vitripennis* having 51 (Fig. S9). A total of 7 GR genes were identified in *A. nilaparvatae*. The 13 Hymenoptera species generally have 20 to 50 IRs, except the 90 of *N. vitripennis*. The 23 IR genes of *A. nilaparvatae* are dispersed in the seven clades (Fig. S10). Similarly, the numbers of candidate SNMP genes identified in genomes of the 13 Hymenoptera species ranged from 10 to 30, with 13 in *A. nilaparvatae*. The candidate SNMP genes were clustered into 10 subgroups (Fig. S11). The SNMPs of *A. nilaparvatae* occur in all subgroups except the Subgroup10. We also found four SNMPs of *A. nilaparvatae* form a monophyly within the Subgroup1, which are likely resulted from two events of gene tandem duplication.

The insect PPK family is a member of DEG/ENAC, and PPKs usually have conserved cysteine-rich domain in its extracellular loop and two transmembrane domains. PPKs involved in mechanosensory functions including water, salt, osmotic potential and pheromones [51–54]. TRP channels in insects are also correlated with mechanosensation [55]. The TRP proteins have six transmembrane helices, the last two of which are located on both sides of the ring that determines ion selectivity. The TRP superfamily plays a key role in the response to photoacoustic chemical temperature and touch external stimuli [18, 56]. We identified less than 10 PPK candidate genes in each of the 13 species, except *N. nitripennis* (17) and *F. arisanus* (14). Interestingly, there is only one PPK gene found in the genome of *A. nilaparvatae*, which is the lowest among all species. The candidate PPK genes of the 13 Hymenoptera species were clustered into 9 subgroups and the only PPK gene of *A. nilaparvatae* is homologous to PPK28 (XP_023245644.4) of *C. floridanum* (Fig. S12). Both *A. mellifera* and *N. vitripennis* have 45 TRPs, and *T. pretiosum* and *B. terebrae* have 36. The other species have fewer TRPs. The 9 TRPs of *A. nilaparvatae* dispersed in the five phylogenetic clades of all TRP candidate genes (Fig. S13).

Discussion

Hymenoptera is the second largest order of insects after Coleoptera. It is estimated that Hymenoptera consists of one million species. However, available genomic resources of Hymenoptera species are not rich. As of December 2021, NCBI had included 228 complete genomes of Hymenoptera species (including 55 Formicidae species), but none of these genomes is from the species of family Mymaridae. Here we sequenced and assembled the first whole genome of Mymaridae.

80% of the sequenced genomes of Hymenoptera species are between 180 and 340 Mb in size, with a few exceptions [57–59]. The size of *A. nilaparvatae* genome assembly is 488.8Mb, relatively large in Chalcidoidea. This large genome is likely attributed to the high content of repeat sequences (45.5%), which is higher than the majority of Chalcidoidea insects. Nonetheless, the content of repeat sequences in the genomes of Hymenoptera species can be very low, such as Apidae (< 10%) [60], or very high, such as *B. terebrae* (80.6%). The large genome size of *A. nilaparvatae* contrasts with its extremely small body size (0.6–0.7mm in body length), suggesting that genome size is not positively correlated with body size. We predicted 16861 genes in the genome of *A. nilaparvatae*, within the range of 12,000–20,000 genes [60] usually found in Hymenoptera species. Notably, the genomes of Hymenoptera species are generally low in GC content, ranging from 30–45% [61], but *A. nilaparvatae* even reduces the lower bound to 27.52%. The low GC content in genomes is likely due to GC biased gene transformation and high recombination rate [62].

Previous studies based on transcriptomes have suggested that Chalcidoidea originated at late Jurassic 129–81Mya, and the earliest divergence within Chalcidoidea emerged in early-middle Cretaceous [4, 6]. Our phylogenomic analyses based on whole genomes agree well with previous assessments, that Mymaridae is sister to other groups of Chalcidoidea. We reduced the range of the estimated age of the divergence between Mymaridae and other Chalcidoidea groups from 89 ~ 208 Mya [6] to 116.8–132.1 Mya. The estimates of the mode are well consistent, which is 129 Mya by Peters et al and 126.9 Mya in our study. Our phylogeny also shows difference to the one in Peters et al. (2018). Peters suggested Trichogrammatidae is a basal lineage (second to Mymaridae) in Chalcidoidea [6], but we found Trichogrammatidae is nested deep in Chalcidoidea, sister to Encyrtidae.

The number of chemoreceptor and mechanoreceptor genes differ among species of Hymenoptera, which is related to the complexity of chemoreceptor and mechanoreceptor in these species. Notably, the limitations of available data and method might have hindered accurate identification of these genes. For example, quite a few previous studies were based on low-quality genome assemblies or transcriptomes [21, 25, 28, 29]. We used the latest high-quality genome assemblies to obtain reliable identification of chemoreceptor and mechanoreceptor genes.

OBP is the first soluble chemoreceptor protein found in insects. In Hymenoptera, the number of OBP genes varies greatly among species [25, 63, 64]. We found an expansion of OBP family in Pteromalidae species, relative to other Hymenoptera species. In addition to Pteromalidae, *T. pretiosum* in Trichogrammatidae also has a large OBP family. The OBP phylogenetic analysis showed long branches in Pteromalidae and *T. pretiosum*, suggesting of rapid evolution. The number of CSP genes show lower level of differentiation among species. The NPC2 genes are usually found with large number in Chelicerata species, but not in insects [39]. Consistently, we identified less than 10 members of NPC2 family in each of the 13 Hymenoptera species. The numbers of OBP, CSP and NPC2 in *A. nilaparvatae* are 17, 11 and 4, respectively, within the lower half among the investigated species. Among all species, *C. solmsi* has the lowest number of the three kinds of soluble protein genes, which may be attributed to its strict specificity to the host figs [31].

All of the 13 Hymenoptera species have a large OR family but a small GR family, both of which belong to the GPCR superfamily [17]. Tandem duplication has been extensively found in the OR family, and the expansion of OR family is usually accompanied with a deficient of GR family [17]. IR genes belong to the iGluR family. These numbers of IRs in the 13 insect species are less than the numbers of ORs, but more than the numbers of GRs. In addition, the numbers of SNMPs show little difference among Hymenoptera species, and we identified more SNMPs than previous efforts [65, 66].

PPKs play important roles in mechanosensory and other functions. PPKs evolved under the genetic birth-and-death model, which produces lineage-specific expansions that form local clusters in phylogeny. Among all insect orders, Hymenoptera has the fewest PPK proteins [28]. We found only one PPK gene in *A. nilaparvatae*, the least in the 13 Hymenoptera insects. Similarly, the number of TRP genes in *A. nilaparvatae* is also the smallest in the 13 species. TRP family is classified into seven subfamilies (TRPC, TRPA, TRPM, TRPL, TRPV, TRPN and TRPP), and TRPC is the “classical TRPs” [67]. The size of the TRP family is generally 13–14 in insect species, and the subfamily TRPP is usually absent in Hymenoptera species [24]. Our phylogenetic of TRPs identified in the 13 species clustered the candidates into five subgroups, with TRPL and TRPV being not distinguishable.

Among all the 13 species, *N. vitripennis* has the most abundant chemo- and Mechano- sensory genes, consistent with the observation of more gene family expansions than contractions in the CAFÉ analysis. *T. pretiosum*, another egg parasitoid, has a large number of PPK and TRP genes, as well as abundant chemoreceptor genes. It is likely due to the wide range and complex habitats of the host of *T. pretiosum*. In contrast, the chemo- and mechano-sensory genes of *A. nilaparvatae* have degraded, with less copies than other species in Hymenoptera. Interestingly, the range of the host of *A. nilaparvatae* is very narrow.

Similarly, *C. solmsi*, which is strictly coevolving with the host ficus, also shows a contraction of gene families. This implies a correlation between the numbers of chemoreceptor and mechanoreceptor genes and the single host and simple habitat.

Conclusions

In this study, we assembled and annotated a high-quality, full-length genome the first time for Mymaridae species. The chemosensory and mechanosensory genes of *A. nilaparvatae* and 12 other Hymenoptera species were analyzed. This work provides not only new genomic sequences for the phylogeny of Hymenoptera, but also novel biological insights into host selection and oviposition behavior of egg parasitoids. As the dominant natural enemy of brown planthopper in rice paddy ecosystems, *A. nilaparvatae* is the key biological factor to control brown planthopper. The availability and utilization of the *A. nilaparvatae* genome resources would provide a basis for further protection and utilization of parasitic natural enemies for pest control.

Materials And Methods

Insects

Individuals of the host, *Nilaparvata lugens*, were collected from rice paddy fields at the farm of the South China Agricultural University in Guangdong Province (N 23°9'3", E 113°20'2") in 2016 and reared with rice hydroponic seedlings. Individuals of *A. nilaparvatae* were collected in the paddy field of the farm of South China Agricultural University in 2018 and were stably cultured for 60 generations on rice seedlings with the eggs of *N. lugens* in an insect cage (120 mesh gauze). The insect cage was placed in an insect incubator (GXZ-380D, Ningbo Jiangnan Instrument Factory, Zhejiang, China), and the rearing conditions were as follows: 14:10 h (L:D) photoperiod, 27°C temperature, and 80% humidity.

DNA extraction

The *A. nilaparvatae* wasps used for DNA extraction are F3 descendants of the same pair of ancestors. The wasps of third generation were separately raised in a transparent glass tube (1ml) to prevent the female wasps from mating with males. After emergence, the females were selected for parthenogenesis, and their offspring were all male. About 600 male individuals of *A. nilaparvatae* were used for DNA extraction. The wasps were frozen and grounded with liquid nitrogen by mortar, and we extracted DNA using the Insect DNA Kit (GBCBio Technologies, Guangzhou, China) according to the manual. DNA concentration and purity were examined by Nanodrop 2000c (Thermo Fisher, Waltham, MA, USA).

Genome assembly and quality control

The genome library was constructed and sequenced in the Tianjin Biochip Corporation (Tianjin, China). The Illumina HiSeq X ten platform (San Diego, CA, USA) was used for sequencing, and a 350 bp (insertion size) double-ended library was constructed by splicing DNA. The Kmer method was used to estimate the genome size and genome heterozygosity. Using SOAPdenovo (<http://soap.genomics.org.cn/>, v1.05) [68], we performed preliminary assembling with multiple trials of different Kmer parameters, and then we used the GapCloser software to perform local gap filling and base correction [69]. We also used the PacBio sequel II platform to *de novo* sequence the genome and we used the HiFiasm software (v0.14) (with the parameters -t32 and -f39) [79] to *de novo* assemble the genome. The Illumina short reads were then aligned to the corrected HiFiasm contigs using BWA-MEM (v0.7.17) [71], and Pilon (v1.2) [72] was used to correct errors in the contigs. Genomic integrity was assessed by mapping the single-copy homologous genes from the BUSCO database and the raw reads of Illumina sequencing to the genome assembly. The information of the assembly was visualized by a circle graph using the Circos software.

Gene Annotation

We used RepeatModeler and RepeatMasker (<http://www.repeatmasker.org>, v4.1.1) [73] to *de novo* predict and mask repeat sequences in the assembly. With repeats masked, we predicted protein-coding genes using a combination of *ab initio*, homology-based and transcriptome-based methods. For the homology-based prediction, the genomes of four relative species *Apis cerana*, *Bombus vancouverensis*, *Drosophila Melanogaster* and *Osmia Bicornis* were used. For the *ab initio* prediction, Augustus (V2.5.5) [74] and GeneMark (V4.32) [75] were used. For the transcriptome-based prediction, the transcriptomes of *A. nilaparvatae* were used [76]. Finally, EVM (Evecemodeler) was used to integrate the predicted candidates from different sources, in which the transcriptome-based prediction was given the highest weight [77]. We compared the integrated annotations with the NR databases and removed the scaffolds with more than half of the genes are non-eukaryotic. We also blasted the final set of predicted genes in the databases of NR, COG, GO and KEGG to obtain the functional descriptions of genes.

Phylogenetic analysis and gene family analyses

Orthofinder v. 2.3.3 [78] was used to cluster the protein-coding genes of *A. nilaparvatae*, *Ceratosolen solmsi*, *Pteromalus puparum*, *Trichomalopsis sarcophagae*, *Nasonia vitripennis*, *Copidosoma floridanum*, *Trichogramma pretisum*, *Trichogramma brassicae*, *Diachasma alloeu*, *Fopius arisanus*, *Microplitis demolitor*, *Belonocnema treatae*, and *Apis mellifera* into orthologous groups [79]. We identified 205 single-copy groups (with one member for each species). For each of the single-copy groups, we aligned the protein sequences using MAFFT and transformed to DNA codons using PAL2NAL v. 14.0 [80]. All the genes were then concatenated to construct a phylogeny using RaxML v8.2.9 with the maximum likelihood method [81]. *A. mellifera* (Apidae: Apidae), which belongs to Aculeata, is the outgroup of this evolutionary tree. All the other species are Terebrantia. The phylogeny tree was dated using the MCMCTREE, which is a part of the software packages PAML v.4.9 [82]. In dating the tree, the root age is in prior set between 203 and 276 Mya (Peters et al. 2018) [6]. The divergence between Chalcidoidea and all the other families was constraint at > 130 Mya [20] and the divergence between *A. nilaparvatae* and other Chalcidoidea species was constraint at 99 ~ 130 Mya [6, 20]. We used CAFE v 3.1 to infer the expansion and contraction of gene families along branches of the phylogeny.

To identify genes related to chemo- and mechano-sensory, we blasted the sequences of OBP, CSP, NPC2, OR, IR, SNMP, GR, PPK and TRP genes of *A. mellifera* against the genomes of all the 13 species using BLAST V2.7.1 +. We checked the conservative domains of the candidates manually in Pfam (<http://pfam.xfam.org/search/sequence>) and removed the candidates without typical elements of the domain of the corresponding gene family. The reliable candidates were aligned using MAFFT v7.453. RaxML V8.2.9 [81] was used to construct gene trees under the optimal substitution model selected by the ProtTest v3.4.2 for each gene family.

Abbreviations

bp
base pair
Mb
million bps
Gb
giga bps
BUSCO
Benchmarking Universal Single-Copy Orthologs
Mya
million years ago
OBPs
odorant binding proteins
CSPs
chemosensory proteins
NPC2
niemann-pick type C2 proteins
ORs
olfactory receptors
GRs
gustatory receptors
IRs
ionotropic receptors
SNMPs
sensory neuron membrane proteins
PPKs
pickpocket receptors
TRP
transient receptor potential
ODEs
odorant-degrading enzymes
Fig.
figure
SMRT
single molecular real time Sequencing
TEs
transposable elements
LINEs
long interspersed nuclear elements
LTRs
long terminal repeat retrotransposons
SINEs
short interspersed nuclear elements
DNA
DeoxyriboNucleic Acid
RNA
Ribonucleic Acid
Cys
cysteine
NR
NCBI non-redundant protein sequences
COG
Cluster of Orthologous Groups of Proteins

GO
Gene Ontology
KEGG
Kyoto Encyclopedia of genes and genomes
KOG
Clusters of orthologous groups for eukaryotic complete genomes
Pfam
protein family
Anil
Anagrus nilaparvatae
Amel
Apis mellifera
Btre
Belonocnema treatae
Cflo
Copidosoma floridanum
Csol
Ceratosolen solmsi
Dall
Diachasma alloeum
Fari
Fopius arisanus
Mdem
Microplitis demolitor
Nvit
Nasonia vitripennis
Ppup
Pteromalus puparum
Tbra
Trichogramma brassicae
Tpre
Trichogramma pretiosum
Tsar
Trichomalopsis sarcophagae.

Declarations

Acknowledgments

We thank Tianjin Biochip Corporation (Tianjin, China) for help in genome sequencing and library construction. This work was supported by the Special Foundation for Key Research Program of Guangzhou (Grant No. 201804020062), the Agricultural Science and Technology Innovative and Promotion Program of Guangdong Province (2021KJ113).

Funding

This work was supported by the Special Foundation for Key Research Program of Guangzhou (Grant No. 201804020062), the Agricultural Science and Technology Innovative and Promotion Program of Guangdong Province (2021KJ113).

Availability of data and materials

The data of *A. nilaparvatae* genome are available in the GenBank/EMBL/DDBJ repository under the accession number PRJNA807850. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAKSXM000000000. The version described in this paper is version JAKSXM010000000.

Authors contributions

Y. M. and Q. Z. conceived and designed the whole project. Z. X. G. performed all experiments and analyses and wrote the first draft of the manuscript. L. Y. W. performed the experiments and helped collect data. T. F. H., B. Y. W. and B. J. T. prepared figures, tables, and supporting information. G. R. Z. helped edit the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study does not include human or vertebrate subjects. Our insect materials don't include any wild species at risk of extinction. No specific permits are required for sample collection in this study. We comply with relevant institutional, national and international guidelines and legislation for insect study.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Jervis MA, Eller J, Harvey JA. Resource acquisition, allocation, and utilization in parasitoid reproductive strategies. *Annu Rev Entomol.* 2008; 53: 361–85.
2. Meyhöfer R, Casas J, Dorn S. Mechano- and chemoreceptors and their possible role in host location behavior of *Sympiesis sericeicornis* (Hymenoptera: Eulophidae). *Annals of the Entomological Society of America.* 1997; 2: 208–19.
3. Pennacchio F, Strand MR. Evolution of developmental strategies in parasitic hymenoptera. *Annual Review of Entomology.* 2006; 51:233–58.
4. Poinar G, Huber JT. A new genus of fossil Mymaridae (Hymenoptera) from Cretaceous amber and key to Cretaceous mymarid genera. *Zookeys.* 2011; 130:461–72.
5. Heraty JM, Burks RA, Cruaud A, Gibson GAP, Liljeblad J, Munro J, et al. A phylogenetic analysis of the megadiverse Chalcidoidea (Hymenoptera). *Cladistics.* 2013; 29(5):466–542.
6. Peters RS, Niehuis O, Gunkel S, Blaser M, Mayer C, Podsiadlowski L, et al. Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary success. *Mol Phylogenet Evol.* 2018; 120:286–96.
7. Chen XX, Tang P, Zeng J, van Achterberg C, He JH. Taxonomy of parasitoid wasps in China: An overview. *Biol Control.* 2014; 68:57–72.
8. Zheng XS, Lu YH, Zhu PY, Zhang FC, Tian J, Xu HX, et al. Use of banker plant system for sustainable management of the most important insect pest in rice fields in China. *Sci Rep-Uk.* 2017; 7.
9. Lou YG, Du MH, Turlings TC, Cheng JA, Shan WF. Exogenous application of jasmonic acid induces volatile emissions in rice and enhances parasitism of *Nilaparvata lugens* eggs by the parasitoid *Anagrus nilaparvatae*. *J Chem Ecol.* 2005; 31(9):1985–2002.
10. Zhu PY, Gurr, GM, Lu Z, Heong K, Chen G, Zheng X, et al. Laboratory screening supports the selection of sesame (*Sesamum indicum*) to enhance *Anagrus spp.* parasitoids (Hymenoptera: Mymaridae) of rice planthoppers. *Biol Control.* 2013; 64(1): 83–9.
11. Li CZ, Sun H, Gao Q, Bian FY, Noman A, Xiao WH, Zhou GX, Lou YG. Host plants alter their volatiles to help a solitary egg parasitoid distinguish habitats with parasitized hosts from those without. *Journa.* 2020; 43(7):1740–50.
12. Jacquín-Joly E, Merlin C. Insect olfactory receptors: contributions of molecular biology to chemical ecology. *J Chem Ecol.* 2004; 30(12):2359–97.
13. Benton R, Vannice KS, Vosshall LB. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature.* 2007; 450(7167):289–93.
14. Leal WS. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol.* 2013; 58:373–81.
15. Clyne PJ, Warr CG, Carlson JR. Candidate taste receptors in *Drosophila*. *Science.* 2000; 287(5459):1830–34.
16. Scott K, Brady R, Jr., Cravchik A, Morozov P, Rzhetsky A, et al. A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*. *Journa.* 2001; 104(5):661–73.
17. Robertson HM, Wanner KW. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family. *Journa.* 2006; 16(11):1395–403.
18. Fowler MA, Montell C. *Drosophila* TRP channels and animal behavior. *Journa.* 2013; 92(8–9):394–403.
19. Masague S, Cano A, Asparch Y, Barrozo RB, Minoli S. Sensory discrimination between aversive salty and bitter tastes in an haematophagous insect. *Journa.* 2020; 51(9):1867–80.
20. Maksoud S, Azar D, Granier B, Geze R. New data on the age of the Lower Cretaceous amber outcrops of Lebanon. *Journa.* 2017; 26(2):331–38.
21. Vieira FG, Foret S, He X, Rozas J, Field LM, Zhou JJ. Unique features of odorant-binding proteins of the parasitoid wasp *Nasonia vitripennis* revealed by genome annotation and comparative analyses. *Journa.* 2012; 7(8): e43034.
22. Robertson HM, Wanner KW. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family. *Journa.* 2006; 16(11):1395–403.
23. Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson TJ, Benton R. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *Journa.* 2010; 6(8).
24. Matsuura H, Sokabe T, Kohno K, Tominaga M, Kadowaki T. Evolutionary conservation and changes in insect TRP channels. *Journa.* 2009; 9:228.
25. Foret S, Maleszka R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Journa.* 2006; 16(11):1404–13.
26. Foret S, Wanner KW, Maleszka R. Chemosensory proteins in the honey bee: Insights from the annotated genome, comparative analyses and expressional profiling. *Journa.* 2007; 37(1):19–28.
27. Nichols Z, Vogt RG. The SNMP/CD36 gene family in Diptera, Hymenoptera and Coleoptera: *Drosophila melanogaster*, *D. pseudoobscura*, *Anopheles gambiae*, *Aedes aegypti*, *Apis mellifera*, and *Tribolium castaneum*. *Journa.* 2008; 38(4):398–415.

28. Latorre-Estivalis JM, Almeida FC, Pontes G, Dopazo H, Barrozo RB, Lorenzo MG. Evolution of the Insect PPK Gene Family. *Journa*. 2021; 13(9).
29. Zhou X, Rokas A, Berger SL, Liebig J, Ray A, Zwiebel LJ. Chemoreceptor Evolution in Hymenoptera and Its Implications for the Evolution of Eusociality. *Journa*. 2015; 7(8):2407–16.
30. Tvedte ES, Walden KKO, McElroy KE, Werren JH, Forbes AA, Hood GR, Logsdon JM, Feder JL, Robertson HM. Genome of the parasitoid wasp *Diachasma alloeum*, an emerging model for ecological speciation and transitions to asexual reproduction. *Journa*. 2019; 11(10):2767–73.
31. Wang N, Wang NX, Niu LM, Bian SN, Xiao JH, Huang DW. Odorant-binding protein (OBP) genes affect host specificity in a fig-pollinator mutualistic system. *Journa*. 2014; 23(5):621–31.
32. Bruno D, Grossi G, Salvia R, Scalar A, Farina D, Grimaldi A, Zhou JJ, Bufo SA, Vogel H, Grosse-Wilder E *et al*. Sensilla morphology and complex expression pattern of odorant binding proteins in the vetch aphid *Megoura viciae* (Hemiptera: Aphididae). *Journa*. 2018; 9.
33. Pelosi P, Iovinella I, Zhu J, Wang G, Dani FR. Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Journa*. 2018; 93(1):184–200.
34. Sandler BH, Nikonova L, Leal WS, Clardy J. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Journa*. 2000; 7(2):143–51.
35. Campanacci V, Lartigue A, Hallberg BM, Jones TA, Giudici-Orticoni MT, Tegoni M, Cambillau C. Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. *Journa*. 2003; 100(9):5069–74.
36. Tegoni M, Campanacci V, Cambillau C. Structural aspects of sexual attraction and chemical communication in insects. *Journa*. 2004; 29(5):257–64.
37. Sanchez-Gracia A, Vieira FG, Rozas J. Molecular evolution of the major chemosensory gene families in insects. *Journa*. 2009; 103(3):208–16.
38. Ishida Y, Tsuchiya W, Fujii T, Fujimoto Z, Miyazawa M, Ishibashi J, Matsuyama S, Ishikawa Y, Yamazaki T. Niemann-Pick type C2 protein mediating chemical communication in the worker ant. *Journa*. 2014; 111(10):3847–52.
39. Pelosi P, Iovinella I, Felicioli A, Dani FR. Soluble proteins of chemical communication: an overview across arthropods. *Journa*. 2014; 5:320.
40. Zheng Y, Wang SN, Peng Y, Lu ZY, Shan S, Yang YQ, Li RJ, Zhang YJ, Guo YY. Functional characterization of a Niemann-Pick type C2 protein in the parasitoid wasp *Microplitis mediator*. *Journa*. 2018; 25(5):765–77.
41. Lundin C, Kall L, Kreher SA, Kapp K, Sonnhammer EL, Carlson JR, von Heijne G, Nilsson I. Membrane topology of the *Drosophila* OR83b odorant receptor. *Journa*. 2007; 581(29):5601–04.
42. Benton R, Vannice KS, Gomez-Diaz C, Vossell LB. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Journa*. 2009; 136(1):149–62.
43. Senthilan PR, Piepenbrock D, Ovezmyradov G, Nadrowski B, Bechstedt S, Pauls S, *et al*. *Drosophila* auditory organ genes and genetic hearing defects. *Journa*. 2012; 150(5):1042–54.
44. Koh TW, He Z, Gorur-Shandilya S, Menuz K, Larter NK, Stewart S, Carlson JR. The *Drosophila* IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. *Journa*. 2014; 83(4):850–65.
45. Knecht ZA, Silbering AF, Cruz J, Yang L, Croset V, Benton R, Garrity PA. Ionotropic Receptor-dependent moist and dry cells control hygrosensation in *Drosophila*. *Journa*. 2017; 6.
46. Vogt RG, Miller NE, Litvack R, Fandino RA, Sparks J, Staples J, Friedman R, Dickens JC. The insect SNMP gene family. *Journa*. 2009; 39(7):448–56.
47. Pregitzer P, Greschista M, Breer H, Krieger J. The sensory neurone membrane protein SNMP1 contributes to the sensitivity of a pheromone detection system. *Journa*. 2014; 23(6):733–42.
48. Febbraio M, Silverstein RL. CD36: implications in cardiovascular disease. *Journa*. 2007; 39(11):2012–30.
49. Rasmussen JT, Berglund L, Rasmussen MS, Petersen TE. Assignment of disulfide bridges in bovine CD36. *Journa*. 1998; 257(2):488–94.
50. Rogers ME, Krieger J, Vogt RG. Antennal SNMPs (sensory neuron membrane proteins) of Lepidoptera define a unique family of invertebrate CD36-like proteins. *Journa*. 2001; 49(1):47–61.
51. Zhong LX, Hwang RY, Tracey WD. Pickpocket is a DEG/ENAC protein required for mechanical nociception in *Drosophila larvae*. *Journa*. 2010; 20(5):429–34.
52. Thistle R, Cameron P, Ghorayshi A, Dennison L, Scott K. Contact chemoreceptors mediate male-male repulsion and male-female attraction during *Drosophila courtship*. *Journa*. 2012; 149(5):1140–51.
53. Pontes G, Pereira MH, Barrozo RB. Salt controls feeding decisions in a blood-sucking insect. *Journa*. 2017; 98:93–100.
54. Matthews BJ, Younger MA, Vossell LB. The ion channel ppk301 controls freshwater egg-laying in the mosquito *Aedes aegypti*. *Journa*. 2019; 8.
55. Guo L, Zhou ZD, Mao F, Fan XY, Liu GY, Huang J, Qiao XM. Identification of potential mechanosensitive ion channels involved in texture discrimination during *Drosophila suzukii* egg-laying behaviour. *Journa*. 2020; 29(5):444–51.
56. Liu C, Montell C. Forcing open TRP channels: Mechanical gating as a unifying activation mechanism. *Journa*. 2015; 460(1):22–25.
57. Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, Magoc T, *et al*. Social evolution. Genomic signatures of evolutionary transitions from solitary to group living. *Journa*. 2015; 348(6239):1139–43.
58. Branstetter MG, Childers AK, Cox-Foster D, Hopper KR, Kapheim KM, Toth AL, *et al*. Genomes of the Hymenoptera. *Journa*. 2018; 25:65–75.
59. Xu H, Ye X, Yang Y, Yang Y, Sun YH, Mei Y, Xiong S, He K, Xu L, Fang Q *et al*. Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps. *Journa*. 2021; 38(12):5539–54.
60. Elsik CG, Worley KC, Bennett AK, Beye M, Camara F, Childers CP, *et al*. Finding the missing honey bee genes: lessons learned from a genome upgrade. *Journa*. 2014; 15:86.

61. Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL. EPIGENETICS STUDIES IN ECOLOGY AND EVOLUTION Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Journa*. 2016; 25(8):1769–84.
62. Kent CF, Minaei S, Harpur BA, Zayed A. Recombination is associated with the evolution of genome structure and worker behavior in honey bees. *Journa*. 2012; 109(44):18012–17.
63. He Y, Wang K, Zeng Y, Guo Z, Zhang Y, Wu Q, et al. Analysis of the antennal transcriptome and odorant-binding protein expression profiles of the parasitoid wasp *Encarsia formosa*. *Journa*. 2020; 112(3):2291–301.
64. Zhou YN, Xie S, Chen JN, Wang ZH, Yang P, Zhou SC, et al. Expression and functional characterization of odorant-binding protein genes in the endoparasitic wasp *Cotesia vestalis*. *Journa*. 2021; 28(5):1354–68.
65. Hu YY, Xu SF, Wubie A, Li W, Huang JX, Guo ZB, Zhou T. Molecular characterization and tissue localization of sensory neuron membrane protein from Chinese honey bee, *Apis cerana cerana* (Hymenoptera: Apidae). *Journa*. 2013; 48(4):533 – 45.
66. Sheng S, Liao CW, Zheng Y, Zhou Y, Xu Y, Song WM, et al. Candidate chemosensory genes identified in the endoparasitoid *Meteorus pulchricornis* (Hymenoptera: Braconidae) by antennal transcriptome analysis. *Journa*. 2017; 22:20–31.
67. Montell C. Drosophila TRP channels. *Journa*. 2005; 451(1):19–28.
68. Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Journa*. 2014; 30(12):1660–6.
69. Xu M, Guo L, Gu S, Wang Q, Zhang R, Peters BA, et al. TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Journa*. 2020; 9(9).
70. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Journa*. 2021; 18(2):170–5.
71. Houtgast EJ, Sima VM, Bertels K, Al-Ars Z. Hardware acceleration of BWA-MEM genomic short read mapping for longer read lengths. *Journa*. 2018; 75:54–64.
72. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *Journa*. 2014; 9(11): e112963.
73. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Journa*. 2009; Chap. 4: Unit 4–10.
74. Stanke M, Morgenstern B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Journa*. 2005; 33(Web Server issue): W465-7.
75. Besemer J, Borodovsky M. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Journa*. 2005; 33(Web Server issue): W451-4.
76. Ma Y, Huang TF, Tang BJ, Wang BY, Wang LY, Liu JB, et al. Transcriptome analysis and molecular characterization of soluble chemical communication proteins in the parasitoid wasp *Anagrus nilaparvatae* (Hymenoptera: Mymaridae). *Ecology and Evolution*. 2022; 12 (3): e8661.
77. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Journa*. 2008; 9(1): R7.
78. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Journa*. 2019; 20(1):238.
79. Mei Y, et al. 2022. InsectBase 2.0: a comprehensive gene resource for Insects. *Nucleic Acids Research* 50: Database issue D1040-D1045. doi: 10.1093/nar/gkab1090
80. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Journa*. 2006; 34(Web Server issue): W609-12.
81. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Journa*. 2014; 30(9):1312–3.
82. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Journa*. 2007; 24(8):1586–91.

Figures

Figure 1

Life Cycle of *Anagrus nilaparvatae*. The *A. nilaparvatae* were reared at 14:10 h (L:D) photoperiod, 27°C temperature. (A) Females and males of *A. nilaparvatae* are attracted to the volatiles of rice or planthoppers and fly to parasitize; (B) A female of *A. nilaparvatae* lays eggs. (C) The late larval of *A. nilaparvatae* is pink. (D) In the late pupa stage, the adult morphology of *A. nilaparvatae* gradually emerges, and the sex of which could be distinguished by antennae. (E) The adult of *A. nilaparvatae* after emergence bites through the egg shell of rice planthopper and fly out.

Figure 2

Genome landscape of the parasitoid wasp *Anagrus nilaparvatae*. The letters and numbers outside the circle represent the scaffold label (scaffold length > 3Mb). From outer to inner circles: the density of single nucleotide variants, the heat map of repeat sequence density, the heat map of gene density and the bar graph of GC content; The innermost line shows the collinear genes within the genome, a line connecting a pair of genes.

Figure 3

Phylogenomic analyses of the parasitoid wasp *Anagrus nilaparvatae* and 12 related species. The maximum-likelihood phylogenetic tree was constructed for *A. nilaparvatae* and 12 other hymenopterans based on genomewide single-copy proteins. *Apis mellifera* was used as outgroup. The black numbers on the nodes indicate divergence times (Mya). The expansion (green) and contraction (red) of gene families are shown on the branches.

Figure 4

Maximum-likelihood tree of OBPs of the parasitoid wasp *Anagrus nilaparvatae* and 12 related species. All gene names are the abbreviation of the species name plus the gene serial number, the gene serial number could be found in NCBI (<https://www.ncbi.nlm.nih.gov/>) or InsectBase 2.0 (<http://v2.insect-genome.com/>); OBPs, Odorant Binding Proteins; Anil, *A. nilaparvatae*; Amel, *Apis mellifera*; Btre, *Belonocnema treatae*; Cflo, *Copidosoma floridanum*; Csol, *Ceratosolen solmsi*; Dall, *Diachasma alloeum*; Fari, *Fopius arisanus*; Mdem, *Microplitis demolitor*; Nvit, *Nasonia vitripennis*; Ppup, *Pteromalus puparum*; Tbra, *Trichogramma brassicae*; Tpre, *Trichogramma pretiosum*; Tsar, *Trichomalopsis sarcophagae*.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarydata.mp4](#)
- [supplementaryfile.docx](#)