Peer-reviewed version available at *Insects* **2020**, *11*, 199; doi:10.3390/insects11030199

Article

Screening of Differentially Expressed Microsporidia Genes from *Nosema ceranae* Infected Honey Bee by Suppression Subtractive Hybridization

Zih-Ting Chang[1¶], Chong-Yu Ko[1], Ming-Ren Yen[2], Yue-Wen Chen[1*] and Yu-Shin Nai [3*¶]

[1] Department of Biotechnology and Animal Science, National Ilan University, No. 1, Sec. 1, Shen Nung Road, Ilan, 26047 Taiwan.
[2] Genomics Research Center, Academia Sinica, No. 128, Academia Road, Sec. 2, Nankang District, Taipei 115, Taiwan.
[3] Department of Entomology, National Chung Hsing University, No. 145, Xingda Road, Taichung, 402 Taiwan.
¶ These authors contributed equally to this work
* Correspondence: chenyw@niu.edu.tw (Y.W. Chen); ysnai@nchu.edu.tw (Y.S. Nai)

Abstract: The microsporidium *Nosema ceranae* is a high prevalent parasite of the European honey bee (*Apis mellifera*). This parasite is spreading across the world into its novel host. The developmental process and some mechanisms of *N. ceranae* infected honey bees has been studied thoroughly, however, few studies have been carried out in the mechanism of gene expression in *N. ceranae* during infection process. We therefore performed suppressive subtractive hybridization (SSH) approach to investigate the candidate genes of *N. ceranae* during its infection process. Each 96 clones of infected (forward) and non-infected (reverse) library were dipped onto the membrane for hybridization. A total of 118 differentially expressed sequence tags (ESTs) had been sequenced. For the host responses, 20% of ESTs (13 ESTs, 10 genes and 1 non-coding RNA) from forward library and 83% of ESTs (44 ESTs, 28 genes) from reverse library were identified as differentially expressed genes (DEGs) of the hosts. High percentage of DEGs involved in catalytic activity and metabolic processes, revealed the host gene expression change after *N. ceranae* infection might lead to the unbalance of physiological mechanism. Among the ESTs from forward library, 75.4% ESTs (49 ESTs belonged to 24 genes) were identified as *N. ceranae* genes. Of 24 *N. ceranae* genes, nine DEGs were subjected to real-time quantitative reverse transcription PCR (real-time qRT-PCR) for validation. The results indicated that these genes showed highly up-regulated during *N. ceranae* infection. Among nine *N. ceranae* genes, one *N. ceranae* gene (AAJ76_1600052943) showed the highest expression level after infection. These identified differentially expressed genes from this SSH could provide information about the pathological effects of *N. ceranae*. Validation of nine up-regulated *N. ceranae* genes revealed highly potential for the detection of early nosemosis in the field and provide insight for further applications.

Keywords: microsporidia; *Nosema ceranae*; honey bee; cDNA subtraction

1. Introduction

Honeybee (*Apis mellifera*) is essential pollinator of many crop plants in natural ecosystems and agricultural crops; it contributes more than $14 billion to agriculture annually [1-4]. However, the American apiculture industry experienced catastrophic losses of unknown origin since 2006. The decline and disappearance of bee species in the natural environment and the collapse of honeybee colonies were defined as colony collapse disorder (CCD) [1]. The consequences of this syndrome are

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

evident as an unexplained disappearance of adult bees, a lack of attention to the brood, reduced colony vigor, and heavy winter mortality without any apparent pathological infection. Reductions of honeybee populations were estimated at 23% over the winter of 2006-2007 [5] and at 36% over the winter of 2007-2008 [6]. Additionally, it leads to reduction of honey production and ultimately to colony collapse[7]. As the matter of fact, the CCD is becoming a worldwide issue and honeybee still sulfuring the threatens from either natural environment (*i.e.*, pesticide residues) or the disease caused by microorganisms (*i.e.*, viruses, bacteria and fungi)

As aforementioned, the honeybee pathogens, such as viruses, microsporidian, mites and parasite, undoubtedly affect the development of honeybee colonies and involved in the CCD. Among these pathogens, microsporidia is one of the most common prevalence in the bee populations [5, 8]. There are two species of honeybee microsporidia including *Nosema apis* and *N. ceranae* [9]. Both of them are obligate intracellular parasitic fungi [10]. The microsporidia are the causation pathogen of nosemosis and usually cause the chronic impacts on honeybee populations [9, 11]. The *N. apis* was first discovered *A. mellifera* in Australia, North America, and Europe and *N. ceranae* was originally found in *A. ceranae* in the 1990s [12, 13]. In 2004-2006, *N. ceranae* infections were detected in *A. mellifera* in Taiwan and European countries [14, 15]. Hereafter, the prevalence of *N. ceranae* was confirmed in the population of *A. mellifera* and it has become world-wide distributed bee pathogens in this decade [16]. *N. ceranae* was coincident with early reports of CCD and has been suggested to be a factor in honey bee declines [5]. It affects adult bees and was recently found in collapsing *A. mellifera* colonies in Spain [17]. Experimental results suggest that *N. ceranae* is more virulent than *N. apis*, therefore, *N. ceranae* is considered to be an important bee pathogen causing bee colony loss [16].

N. ceranae is a gut-pathogen, in terms of the infection of *N. ceranae* usually caused cytopathic effects (CPE) in the infected midgut tissue of honeybees [11]. It has reported that *N. ceranae* infection led to the reduction of honey production, malnutrition, shorter life span, and increasing the mortality of adult honeybees [7, 11, 18-20]. Moreover, the infection of *N. ceranae* would change the metabolism of carbohydrates and suppress metabolism related genes expression [21, 22]. However, beekeepers are not easy to note *N. ceranae* infection in the honeybee colonies due to the lack of obvious symptoms of latent infection of *N. ceranae*. Ignores of latent infection of *N. ceranae* might raise the risk of long-term colony infection of *N. ceranae*, and then cause colony disease outbreaks [16]. Therefore, early detections of nosemosis at the latent infection stage is needed for management of honeybee colonies.

In this study, we attempted to identify the differential expressed genes between *N. ceranae* infection and non-infection honey bees and to clarify the gene expression profile of highly up-regulated gene during *N. ceranae* infection. To this aim, a method of suppression subtractive hybridization (SSH), which is a powerful method for selectively amplifying differentially expressed genes [23], was used for identification of the differential expression *N. ceranae* genes and also the host response genes. The highly up-regulated *N. ceranae* specific genes were further validated and investigated. These data could provide an insight for the detection of early nosemosis in the field.

2. Materials and Methods

2.1 Purification of microsporidia from infected honeybee midguts

Spores for the inoculation were isolated from live heavily infected bees of *Apis mellifera*, collected from a naturally infected hive located in the experimental apiary of the campus of National Taiwan University in Taipei, Taiwan and National Ilan University in Illan, Taiwan. After dissecting the intestinal tracts, the midgut tissues were macerated in distilled water using a manual tissue grinder and examined under phase-contrast microscopy (Olympus IX71). Infected tissues were homogenized and filtered through three layers of stainless net and centrifuged at 3000 ×*g* for 15 min. The pellet was then transferred into aseptic 90% Percoll and 10% 1.5M NaCl followed by centrifugation at 15,000 ×*g* for 40 min. The band centrifuged to the bottom of the tube contained mature spores [24, 25]. The purified spores were stored in distilled water at 4℃. The spore concentration was determined by counting with a haemocytometer chamber and the suspension was freshly prepared before using.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

2.2 Experimental infection

The experimental infection was modified from Higes et al. (2007), briefly described as below, frames of sealed brood obtained from a healthy colony of *A. mellifera* (Nosema free confirmed by PCR, weakly) were kept in an incubator at 32 ± 2℃ and 95% relative humidity to provide newly emerged Nosema free honeybees. Three days after eclosion, the bees were starved for 2 h and 100 bees per group were each fed with 5 µl of 50% sucrose solution containing 1.25×10^5 spores of *N. ceranae* by a droplet with the spore solution at the tip of a micropipette until it had consumed the entire droplet [26]. Control honeybees were fed with 5 µl of 50% sucrose solution. The honey bees were reared in an incubator at 32 ± 2℃ and 95% relative humidity and the survival individuals were recorded daily for 21 days. The Kaplan-Meier survival curve was generated by SPSS software. For the sample collections, 10 bees from each cage were collected at 7, 14 and 21 days post infection (dpi) and their ventriculi processed for microscopic examination. Spores were counted and spore production was recorded every week until 3 weeks pi. Control bees were also sacrificed in day 7, 14 and 21 dpi and analyzed to confirm the absence of spores in ventriculus.

2.3 Detection of infection

DNA was extracted from the midgut tissues of control and infected honey bees by using the following procedure modified from Tsai et al., 2002 [27]. In brief, 10 midgut tissues of control and infected honey bees were homogenized in TE buffer (0.1 M Tris, 0.01 M EDTA, pH 9.0) with a pestle. The macerated solution was centrifuged and the supernatant was discarded. The pellets were incubated with RNase A at 37℃ for 1 hr and proteinase K at 56℃ overnight followed by DNA extraction with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1), and DNA precipitation with 1/10 volume of 3M sodium acetate and 2 times volume of pure ethanol. The DNA pellet was air dried and then dissolved in water. The SSU rDNA fragment of the microsporidium was amplified using primer set 18f /1537r (Supplementary Table 1). For the internal control, 18S primer set 143F/145R (Lo et al., 1994) was used (Supplementary Table 1). Each 50-µl PCR mix contained 5 µl 10× reaction buffer (Bioman), 4 µl 2.5 mM dNTPs, 0.5 µl 100 mM of each primer, 1µl 1.25 U HiFi *Taq* polymerase (RBC) and 1 µl template DNA. PCR amplifications were performed as follows on an AG9600 Themal Station (Biotronics Corp.): thermal cycler was preheated at 95℃ for 5 min, 35 cycles at 94℃ for 30 sec, 50℃ for 1 min, and 72℃ for 2 min, followed by a 10 min final extension at 72℃ and storage at 20℃. The PCR product was cloned into T&A cloning vector (RBC Bioscience) and commercially sequenced (Genomics Biosci. & Tech. Company).

2.4 Construction of subtractive cDNA library

Total RNA was extracted using TRIzol® Reagent (ambion) according to the manufacturer's instructions from 10 of control and infected tissues at 7, 14 and 21 dpi and then pooled together equal quantities of three RNA samples per group to maximize detect differential expression of genes that may show variation between individuals in the exact timing of expression. Single-stranded and double-stranded cDNA were synthesized from the infected and control RNA pools with the SMART™ PCR cDNA Synthesis Kit (Clontech, USA) according to the manufacturer's protocol.

Two subtractive cDNA libraries (forward and reverse subtractive cDNA libraries) from the control and infected honeybees, were constructed by using the PCR-Select cDNA Subtraction kit (Clontech) according to the manufacturer's specifications. Briefly descript below, cDNA from the infected honeybees was used as the tester (forward subtractive cDNA library, fscl.)/driver (reverse subtractive cDNA library, rscl.), and cDNA from the control honeybees as the driver (fscl.)/tester (rscl.). Both tester and driver cDNAs were digested with *Rsa*I to produce shorter bluntended fragments. After digestion with *Rsa*I, the tester cDNA was divided into two portions, each of which was ligated with a different adapter (Adaptor-1 and Adaptor-2R) at 16℃ for overnight. After ligation, each tester cDNA was separately hybridized at 68℃ for 8 h with an excess of driver cDNA after denaturation at 98℃ for 90 sec. Then the two hybridized samples were mixed together and hybridized at 68℃ for overnight with excess of denatured driver cDNA. The resulting mixture was added with 200 ml dilution buffer and amplified by two rounds of suppression PCR. Before the

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

primary PCR, the reaction mixture was incubated at 75℃ for 5 min to extend the adaptors. Primary PCR was performed at 94℃ for 30 sec, 66℃ for 30 sec, and 72℃ for 90 sec for 27 cycles in a reaction volume of 25μl. The PCR product was then diluted 10-fold, and 1 μl-diluted product was used as template in the next subsequently nested PCR. The subtracted secondary PCR products were purified with DNA Clean/Extraction Kit (GeneMark) and then ligated into T&A cloning vector (RBC) followed by transformation into *Escherichia coli*, DH5α competent cells (RBC) to generate subtracted cDNA libraries. Each step of SSH was tested following the manufacturer's instructions. Total 192 white colonies from forward (96 colonies) and reverse libraries (96 colonies) were checked by colonies PCR. The PCR products were further subjected to dot-blotting method, which described from the manufacturer's protocol, and the potential colonies were identified and commercially sequenced (Genomics Bioscience & Techology). DNA sequencing from the 3' end and 5' end of the cDNA was conducted with M13 forward or reverse primers, respectively, on a high throughput automated sequencer (MJ Research BaseStation and ABI3730, USA) using standard protocols.

2.5 DNA sequences analysis

cDNA sequences from the each SSH libraries were sequenced. CodonCode Aligner was used for base-calling (Q > 13), trimming vector sequences, sequences assemblage. All the sequences were filtered for size (less than 100 bp). Putative functions of the unique sequences were discovered by using BLASTn and BLASTp to translate each nucleotide query sequence into all reading frames and then searching for matches in the NCBI non-redundant database. Significant hits (with *E* value < 10^{-10}) in the NCBI nr database were further submitted to the Gene Ontology (GO) enrichment analysis (http://geneontology.org/page/go-enrichment-analysis) for protein functions. The Gene GO annotations that classify proteins by molecular function, biological process and cellular component. Each unique sequence was tentatively assigned GO classification based on annotation of the single "best hit" match. These data were then used to classify the corresponding genes according to their GO functions.

2.6 RT-qPCR validations of microsporidia specific genes

To further validate the differential expressed microsporidia specific genes in *N. cereana* infected honey bees, microsporidia genes have 2 ≧ ESTs present in the forward library and dot-blotting difference ≧ 3-fold were selected and subjected to quantitative RT-qPCR analysis. Gene specific primer sets for RT-qPCR were designed by Primer Express v3.0 (Supplementary Table 1). Total RNA was extracted using TRIzol® Reagent (ambion) according to the manufacturer's instructions from 5 of control and infected midguts at 7, 14 and 21 dpi The extracted RNA (2μg) was treated with DNaseI (Roche Molecular Biochemicals), recovered by a phenol/chloroform / isoamylalcohol extraction and precipitated with ethanol. The RNA was then treated with DNase I (Invitrogen Life Technologies) following the manufacturer's instructions to reduce genomic DNA contamination before RT-qPCR. The DNase I-treated total RNA samples were conducted to the reverse-transcription by GScript RTase kit (GeneDireX, US) following the manufacturer's instructions. The reaction mixtures at 42°C for 1.5 h, were incubated at 42°C for 1.5 h and then the reaction was stopped at 70°C for 15 min. Real-time qPCR was performed by using Thermo Scientific Verso SYBR Green 1-step qRT-PCR ROX Mix kit (Thermo-Fisher) in a 96-well Bio-Rad CFX96 Real-Time PCR System (Bio-Rad). All samples were performed in triplicate. The relative gene expression levels were calculated by $2^{-\triangle\triangle Ct}$ method [28].

3. Results and Discussion

3.1 Microsporidian infection

The midgut tissues of control and infected honey bees were homogenized and mature spores in the macerated solution was counted under microscopy. The formation of mature spores was observed after 14 to 21 dpi, moreover, the spore number increased from 14 dpi. to 21 dpi (Figure 1A and B). The DNA of honeybee midgets was extracted for confirming the infection of *N. ceranae* by PCR with 18f and 1537r primer sets (microsporidia degenerated primer set). The results of PCR

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

revealed that only the infected honeybees were detected (Figure 1B). The infection of *N. ceranae* was detected at 7 dpi though the mature spores were only observed at 14 dpi. (Figure 1). This result was similar to the previous report that different parasite stages of *N. ceranae* should present in the midgut epithelial cells after the experimental infection of *N. ceranae* with honey bees at 6-7 dpi [11], indicated the honey bees were under *N. ceranae* infection. It has been described that the pathological effects of *N. ceranae* infected *A. mellifera* resulted in the reduction of the lifespan [11, 21]. From our survival rate analysis, adult honey bees infected with *N. ceranae* showed lower survival possibility and survival time than those of uninfected honey bees (Figure 1D); this data proved a consistent negative effect of the *N. ceranae* to adult bees.

3.2 Screening of the subtracted cDNA libraries

There are two subtracted cDNA libraries was established, and total of 192 white colonies (each of 96 colonies) were first screened by colony PCR (Figure 2). The results of colony PCR showed that the positive rate was 97.9% for fscl. and 95.8% for rscl. The size of inserted cDNA fragments varied from 180 bp to 1 kb (Figure 2). The positive PCR products were then subjected to dot blotting for screening the differential ESTs. A total of 118 differentially expressed sequence tags (ESTs) were identified in the subtracted libraries and sequenced. There are 65 and 53 ESTs were identified from infected (forward) and non-infected (reverse) subtracted libraries, respectively (Table 1; Figure 2B-C and E-F). These clones revealed the genes with highly differentially expression levels and were further sequenced.

3.3 Analysis of EST sequences

Total of 118 ESTs were sequenced and further analyzed. The average sequence length of ESTs in each library was 426 bp for the infected (forward) library and 375 bp for the non-infected (reverse) library (Table 1). According to the NCBI BLASTn analysis,75.4% ESTs (49 ESTs) were identified as microsporidia genes and 20% ESTs (13 ESTs) were belonged to *Apis* spp.in the fscl (Table 2; Figure 2B-C). From the rscl, there were 83% ESTs (44 ESTs) showed significant homology to insect host, especially Heminoptera insects (*Apis mellifera* and *A. florea*) (Table 3; Figure 2E-F). This result indicated that the microsporidia genes were highly up-regulated during the infection process. There were only few host genes showed the dramatically responses of up-regulated during *N. ceranae* infection, while most of the up-regulated host genes were identified from the rscl, suggested the suppression of host gene expression by *N. ceranae* infection.

The result of Gene Ontology (GO) analysis showed total of 30 DEGs (78.9%) from host were identified and were further categorized into molecular function, biological process and cellular component (Figure 3; Supplementary Table 2). Moreover, there were 16 GO terms belonged to three major categories (Figure 3; Supplementary Table 2).

It has been reported that *N. ceranae* infection could changes in metabolism of host, besides, the *N. ceranae* infection also showed the negative effects on the immune system to honey bees; the honey bees infected with *N. ceranae* showed a down-regulation of some immune-related genes, such as *abaecin*, *apidecin*, *defensin*, *hymenoptaecin*, *glucose dehydrogenase* (*GLD*) and *vitellogenin* (*Vg*), suggesting that *N. ceranae* infection suppresses immune defense mechanisms in honey bees [29-31].

Based on the GO analysis, the genes involved in catalytic activity (GO:0003824) were highly expressed and followed by cellular process (GO:0009987) / membrane (GO:0016020) and binding (GO:0005488) / transporter activity (GO:0005215) / cell (GO:0005623) in the fscl. For the rscl, host genes involved in the metabolic process (GO:0008152) and cell (GO:0005623) showed highly changed during *N. ceranae* infection and followed by catalytic activity (GO:0003824) / binding (GO:0005488) / cellular process (GO:0009987) / membrane (GO:0016020) and transporter activity (GO:0005215) / organelle (GO:0043226) (Figure 3; Supplementary Table 2). In conclusion, according to our data, high percentage of host DEGs involved in catalytic activity (GO:0003824), cellular process (GO:0009987), membrane (GO:0016020) / cell (GO:0005623) and transporter activity (GO:0005215) (Figure 3; Supplementary Table 2). As the matter of fact that the up-regulation of the α-glucosidase gene and genes, which involved in trehalose transport, and the down-regulation of the trehalase and the

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

glucose-methanol-choline oxidoreductase three encoding genes during *N. ceranae* infection. The unbalance of the gene expressions results in the modifications of carbohydrate metabolism in the infected honey bees; suggested this changes in the gene expressions and carbohydrate metabolism were manipulative activity of the *N. ceranae* to get nutrients from host [21, 29]. Based on our data, the genes involved in the metabolic process (GO:0008152) from the rscl showed differential expressed, revealed the host gene expression change after *N. ceranae* infection and lead to the unbalance of physiological mechanism.

3.4 Identification and validation of highly expressed N. cerana specific genes

From the forward (infected) library, there are 75.4% ESTs (49 ESTs belonged to 24 genes) were identified as *N. ceranae* genes (Table 2). Of 24 *N. ceranae* genes, nine DEGs, which have more than 2 identified ESTs and >3-fold change by dot bolt analysis were selected for further validation by qRT-PCR, including four hypothetical proteins (AAJ76_1600052943, AAJ76_500092411, AAJ76_2000141845 and AAJ76_1900028347), *14-3-3 protein 1-like protein* (AAJ76_300017093), *Heat shock protein 90* (AAJ76_1170002375), *Ubiquitin* (AAJ76_1300017036), *Forkhead hnf3 transcription factor* (AAJ76_500091146) and *Histone H3* (AAJ76_800037613) (Figure 4). Similar to the expectations, the results indicated that these genes showed highly up-regulated during *N. ceranae* infection. Moreover, the results also showed that all microsporidia specific gene expressions had the highest gene expression level at 7 dpi, and decreased at the 14 and 21 dpi (Figure 5A and B). It should be noted that among these 9 *N. ceranae* specific genes, a *hypothetical protein* (AAJ76_1600052943) (the previous library clone no.= sr22 and provisional named sr22 gene) has the consistently the highest expression, which up-regulated to ~1.1×10^4 fold at 7 dpi. to the mock infection 2340 and 1148 fold at 14 and 21 dpi, respectively (Figure 5B). For the other eight genes, there are three genes, including *histone H3*, *ubiquitin*, and *hypothetical protein* (AAJ76_2000141845) showed higher expression levels, which ranged from ~2100 to 4400 fold to the mock infection at 7 dpi. and the *hypothetical protein* (AAJ76_2000141845) also showed consistently highly expressed at 14 to 21 dpi (Figure 5A).

Ubiquitin is a small protein, which involved in the ubiquitin-proteasome system (UPS) The UPS is the crucial intracellular protein degradation system in eukaryotes. It has been reported that the infection of *N. bombycis* resulted in up-regulation of protein degradation-related enzyme in the host midgut, including the *S-phase kinase-associated protein 1* and the ubiquitin-conjugating enzyme *E2G*. Both of which are involved in the ubiquitin-proteasome pathway [32]. From our data, the highly expression of *ubiquitin* was observed in *N. cerana*, indicated that the host defense mechanism to the parasite and therefore trigger the protein degradation-related enzyme in *N. cerana*.

Besides, the highly expression of *histone H3* was detected in the *N. ceranae* infection process; histone proteins are essential for DNA packing, chromosome stabilization and gene expression in the nucleus of a eukaryotes cell. Each the nucleosome consists of two copies of each core histone H2A, H2B, H3 and H4 to be a octameric protein complex. Histones H3 and H4 are the most highly conserved histones, which play a regulatory role in chromatin formation such as retain the ability to impede transcription [33]. Therefore, it was assumed that highly expression of *N. cerana histone H3* might enhance the regulation of the chromatin formation and facilitate the microsporidia DNA replication.

For the other up-regulated *N. ceranae* genes, they showed up-regulated expression profile during *N. ceranae* infection (Figure 5A). The functions of several genes revealed the response of *N. ceranae* to the honey bee host and also showed the replication activity of *N. ceranae* during infection. For the *14-3-3 protein 1-like protein* (AAJ76_300017093), it is a serine/threonine binding protein. The *14-3-3 protein 1-like protein* inhibits apoptosis through sequestration of pro-apoptotic client proteins [34]. Interestingly, the highly expression levels of *ubiquitin* of *N. cerana* was detected at 7 dpi. and then decreased at 14 and 21 dpi, while the expression of *14-3-3 protein 1-like protein* was higher than that of *ubiquitin* at 21 dpi; besides, the *heat shock protein 90 (Hsp90)* was also identified as highly expressed *N. ceranae* gene that facilitates metastable protein maturation, stabilization of aggregation-prone proteins, quality control of misfolded proteins and assists in keeping proteins in activation-competent conformations [35]; it seems that during the infection process, *N. ceranae* not only needs to against the

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

stress from the host, but also attempted to stabilize the physical pathway itself for the self-propagation purpose. This hypothesis is worthy to be addressed in the future.

For the gene *forkhead hnf3 transcription factor* (AAJ76_500091146), it presented throughout the animal kingdom as well as in fungi and yeast (Mazet et al. 2003). The function of *forkhead hnf3 transcription factor* is characterized by a type of DNA-binding domain known as the forkhead box (FOX) [36]. It has also described that *forkhead hnf3 transcription factor* involved in a wide range of biological functions, including development, growth, stress resistance, apoptosis, cell cycle, immunity, metabolism, reproduction and ageing [37]. As the matter of fact that, in terms of these genes, they might play important roles for the replications of microsporidia, but it still need to be further evaluated. The genome of honeybee (*A. mellifera*) had been sequenced and it consisted of a little more than 10,000 genes, lower than other insects (*D. melanogaster*, 13,600 genes; *A. gambiae*, 14,000 genes; *B. mori*, 18,500 genes) [38]. In additionally, the draft assembly (7.86 MB) and highly compact (~1 gene/kb) genome of the *N. ceranae* genome has been fully sequenced in 2009 [39]. *N. ceranae* has a strongly AT-biased genome (74% A+T) and a diversity of repetitive elements, complicating the assembly. Total 2,614 predicted protein-coding sequences were predicted and 1,366 of these genes have homologs in the microsporidian *Encephalitozoon cuniculi*, the most closely related published genome sequence. These databases could be a reference for studying and clarifying the virulent of *N. ceranae* and the interaction between their host, *A. mellifera* in the near future. From our data, several *N. ceranae* genes were identified as differentially expressed genes through the SSH, which could provide information about the pathological effects of *N. ceranae*. In terms the validation of nine up-regulated *N. ceranae* genes, it revealed highly potential for the detection of early nosemosis in the field and provide insight for further applications.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

3.2. Figures, Tables and Schemes

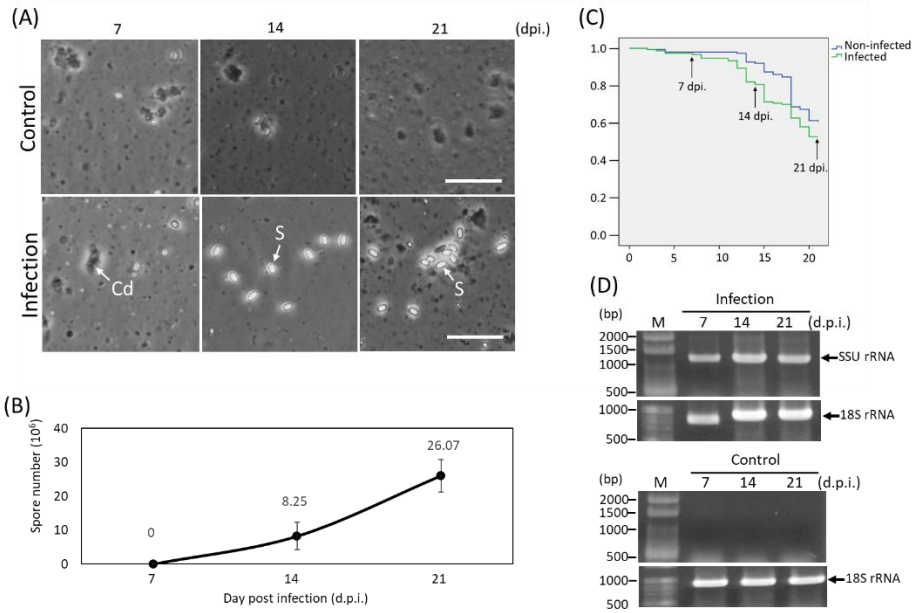

Figure 1. Observations and molecular detection of microsporidia infection at 7, 14 and 21 days post infection (dpi). (A) Light microscope observations; at 14 and 21 dpi, the mature spores (S) were formed. Cd= Cell debris; Scale bar= 25μm; (B) Spore number was counted at each time point; (C) Kaplan-Meier survival curve; (D) Molecular detection of microsporidia infection; SSU rRNA= small subunit ribosomal RNA;18S rDNA= DNA internal control. Control= non-infected honey Infection= infected honey bee.

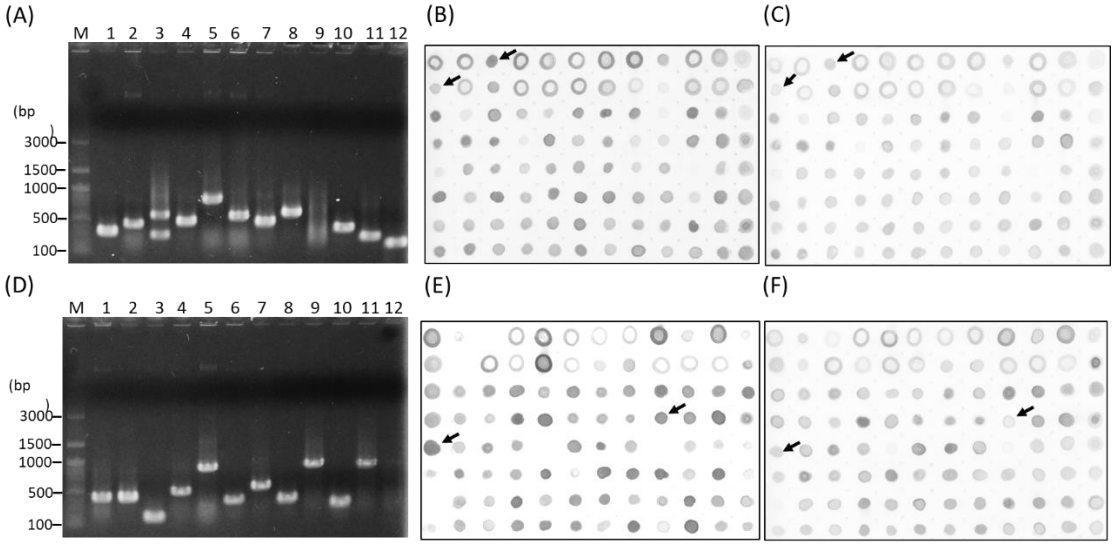

Figure 2. Screening and identification of differential expressed sequence tags (ESTs) from *Nosema ceranae* infected honey bee by suppression subtractive hybridization (SSH) and Dot-blotting. ESTs insert size of (A) forward and (D) reverse SSH cDNA libraries were confirmed by colony PCR. (B) Forward subtractive library (fsl, differential ESTs in infected group) hybridized with forward-subtracted probe and (C) with reverse-subtracted probe. (E) Reverse subtractive library (rsl, differential ESTs in control group) hybridized with reverse-subtracted probe and (F) with forward-subtracted probe. M= 100bp DNA ladder. The arrows indicated the differential ESTs for sequencing.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

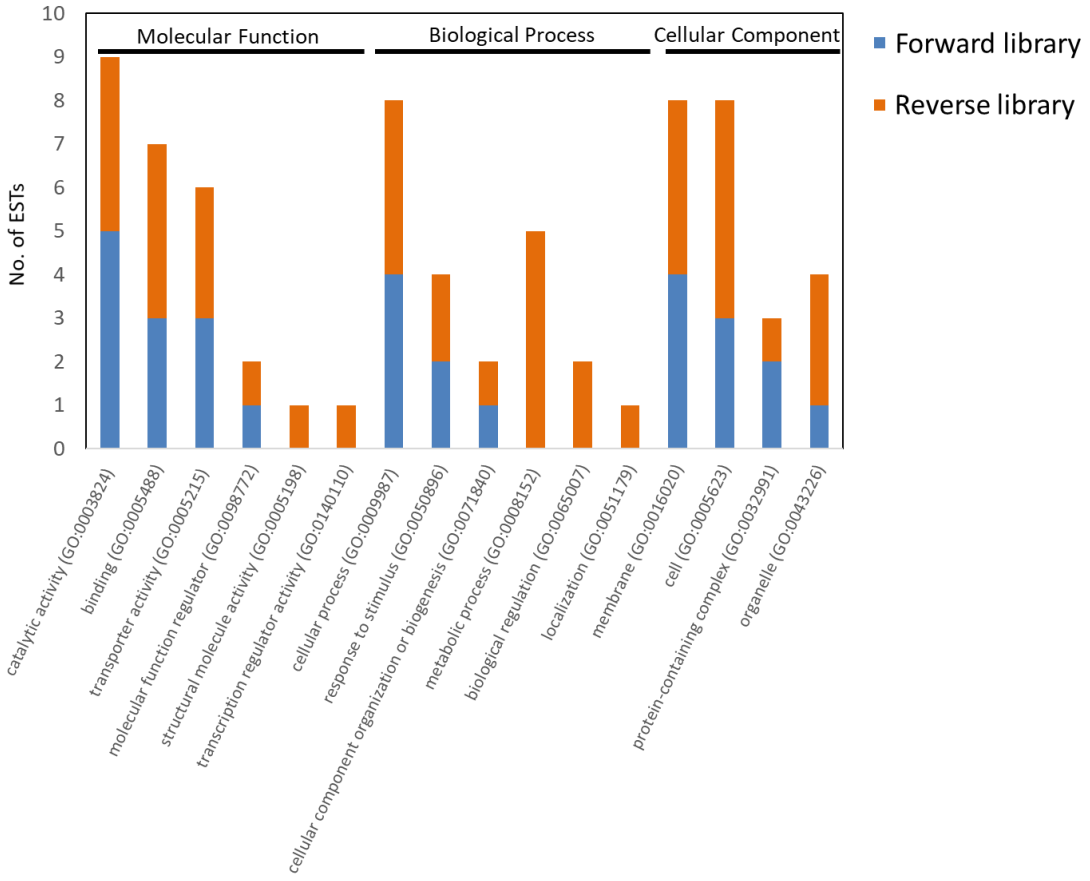

Figure 3. Gene Ontology (GO) analysis of 47 host insect ESTs from forward library (13 ESTs) and reverse library (34 ESTs).

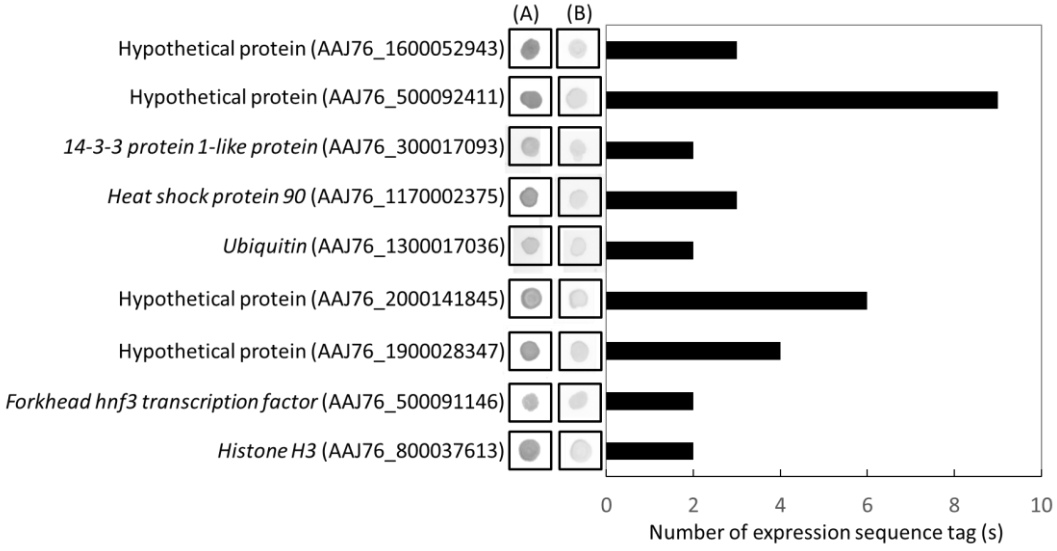

Figure 4. Distribution of expression sequence tags (ESTs) belong to eighteen *N. ceranae* specific genes and dot-blot hybridization test the cDNA library of *N. ceranae*-infected cDNA library. (A) The probes that came from subtractive PCR products of the *N. ceranae*-infected cDNA library was used to test the cDNA library of *N.ceranae*-infected cDNA library; (B) The probes that came from unsubtractive PCR products of control honey bee cDNA library was used to test the cDNA library of *N. ceranae*-infected cDNA library.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

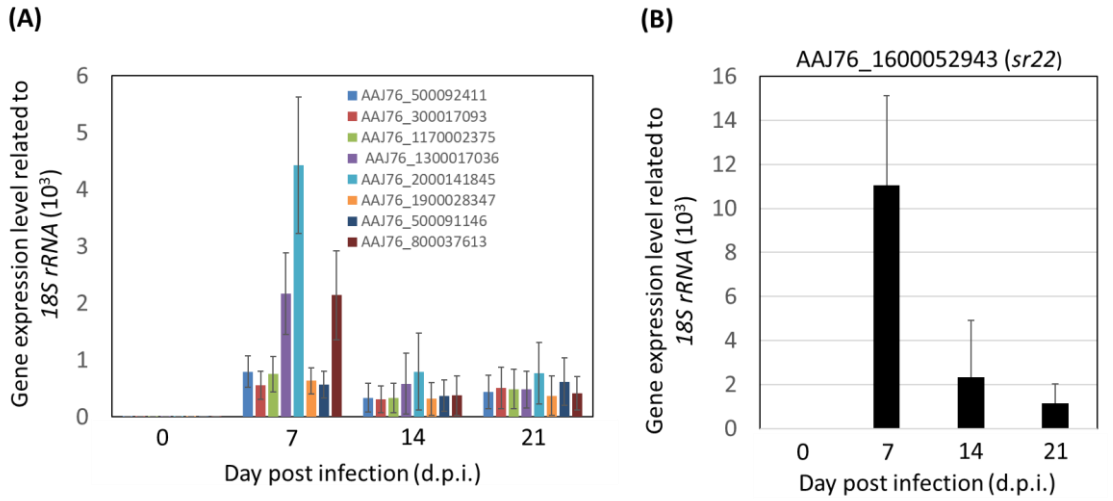

Figure 5. Validation of 9 microsporidia specific gene expression by RT-qPCR at 7, 14 and 21 dpi (A) Expression levels of 8 *N. ceranae* genes after honey bees infected with *N. ceranae* at different days post infections; (B) The highest expressed *N. ceranae* gene, AAJ76_1600052943 (*sr22*).

Table 1. Statistics for the expression sequence tags (ESTs)

Statistic		Value
Total ESTs[a]		118
	Forward library	65
	Reverse library	53
Mean EST sequence length	Forward library	426 bp
	Reverse library	375 bp

[a]ESTs= Expression sequence tags

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

1

Table 2. Analysis of 65 differential expressed sequence tags (ESTs) identified from forward cDNA library

No	EST number	Insert Sequence length (bp)	NCBI_BLAST_N			NCBI_BLAST_X		
			Species	Gene name	E-value	Species	Gene name	E-value
1	SSH-Forward-lib-1	302	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	3.00E-127	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	9.00E-55
2	SSH-Forward-lib-2	396	*Nosema ceranae*	Dynein light chain 1 (AAJ76_5000122115)	4.00E-77	*Nosema ceranae*	Dynein light chain 1	2.00E-24
3	SSH-Forward-lib-3	302	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	2.00E-124	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	5.00E-55
4	SSH-Forward-lib-4	421	*Nosema ceranae*	Hypothetical protein (AAJ76_700047245)	1.00E-177	*Nosema ceranae*	Hypothetical protein NCER_101230 (AAJ76_700047245)	2.00E-64
5	SSH-Forward-lib-5	958	*Nosema ceranae*	60s ribosomal protein L10a (AAJ76_1200039265)	0.00E+00	*Nosema ceranae*	60s ribosomal protein L10a	2.00E-145
6	SSH-Forward-lib-7	413	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	8.00E-86
7	SSH-Forward-lib-8	532	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	1.00E-82
8	SSH-Forward-lib-9	346	*Nosema ceranae*	14-3-3 protein 1-like protein (AAJ76_300017093)	8.00E-158	*Nosema ceranae*	14-3-3 protein 1-like protein	4.00E-66
9	SSH-Forward-lib-10	354	*Nosema ceranae*	Heat shock protein 90 (AAJ76_1170002375)	2.00E-153	*Nosema ceranae*	Heat shock protein 90	7.00E-61
10	SSH-Forward-lib-11	262	*Apis mellifera*	PREDICTED: *Apis mellifera* protein G12-like (LOC102654405)	1.00E-100	*Apis mellifera*	PREDICTED: protein G12-like	1.00E-18
11	SSH-Forward-lib-12	181	*Apis mellifera*	*Apis mellifera* ribosomal protein LP1 (RpLP1 (NM_001185144)	8.00E-63	*Apis mellifera*	Unknown	6.00E-17
12	SSH-Forward-lib-13	326	*Nosema ceranae*	Ubiquitin (AAJ76_1300017036)	1.00E-115	*Nosema ceranae*	Ubiquitin	1.00E-47
13	SSH-Forward-lib-14	231	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	1.00E-88	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	8.00E-34

Peer-reviewed version available at *Insects* **2020**, *11*, 199; doi:10.3390/insects11030199

14	SSH-Forward-lib-15	212	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	3.00E-93	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	2.00E-33
15	SSH-Forward-lib-19	959	*Apis mellifera*	PREDICTED: *Apis mellifera* insulin receptor B (IR-B), transcript variant 1	0.00E+00	*Bacteroides dorei*	Hypothetical protein HMPREF1063_05178	1.00E-18
16	SSH-Forward-lib-20	163	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	1.00E-47	*Nosema ceranae*	Actin	3.00E-16
17	SSH-Forward-lib-21	814	*Schistosoma rodhaini*	Schistosoma rodhaini genome assembly S_rodhaini_Burundi (LL960759)	3.00E+00	-	-	-
18	SSH-Forward-lib-23	231	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	2.00E-93	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	6.00E-35
19	SSH-Forward-lib-25	319	*Nosema ceranae*	Mitochondrial sulfhydryl oxidase (AAJ76_600077605)	3.00E-136	*Nosema ceranae*	Mitochondrial sulfhydryl oxidase	9.00E-58
20	SSH-Forward-lib-26	133	*Alexandrium tamarense*	*Alexandrium tamarense* mRNA (AB233356)	4.00E-14	*Durinskia baltica*	Cytochrome oxidase subunit 3, partial	2.00E-32
21	SSH-Forward-lib-27	537	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	0.00E+00	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	2.00E-116
22	SSH-Forward-lib-29	186	*Apis florea*	PREDICTED: *Apis florea* V-type proton ATPase 16 kDa proteolipid subunit (LOC100867892)	5.00E-66	*Melipona quadrifasciata*	V-type proton ATPase 16 kDa proteolipid subunit, partial	2.00E-32
23	SSH-Forward-lib-30	244	*Nosema ceranae*	Hypothetical protein (AAJ76_6800017741)	4.00E-106	*Nosema ceranae*	Hypothetical protein (AAJ76_6800017741)	9.00E-40
24	SSH-Forward-lib-32	323	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	3.00E-135	*Nosema ceranae*	Hypothetical protein (AAJ76_1600052943)	1.00E-56
25	SSH-Forward-lib-34	885	*Apis mellifera*	PREDICTED: *Apis mellifera* uncharacterized LOC724199 (LOC724199)	0.00E+00	*Apis dorsata*	PREDICTED: extensin-like	2.00E-27
26	SSH-Forward-lib-36	320	*Nosema ceranae*	Polar tube protein 1 (PTP1) gene	2.00E-88	*Vibrio parahaemolyticus*	V-type ATPase subunit C	7.00E-14
27	SSH-Forward-lib-37	232	*Nosema ceranae*	Polar tube protein 2 (AAJ76_1900025375)	3.00E-85	*Nosema ceranae*	Polar tube protein 2	2.00E-34

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

28	SSH-Forward-lib-41	594	*Nosema ceranae*	Heat shock protein 90 (AAJ76_1170002375)	0.00E+00	*Nosema ceranae*	Heat shock protein 90	6.00E-122
29	SSH-Forward-lib-44	682	*Apis mellifera*	PREDICTED: *Apis mellifera* uncharacterized LOC100578731 (LOC100578731)	0.00E+00	*Apis mellifera*	PREDICTED: uncharacterized protein LOC100578731 isoform X1	6.00E-123
30	SSH-Forward-lib-46	411	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	1.00E-87
31	SSH-Forward-lib-48	294	*Nosema ceranae*	Hypothetical protein (AAJ76_1900028347)	2.00E-129	*Nosema ceranae*	Hypothetical protein AAJ76_1900028347	5.00E-51
32	SSH-Forward-lib-49	85	*Apis dorsata*	PREDICTED: *Apis dorsata* V-type proton ATPase 16 kDa proteolipid subunit-like (LOC102679198)	3.00E-27	*Cyprinus carpio*	Hypothetical protein cypCar_00031997	6.00E-06
33	SSH-Forward-lib-51	231	*Nosema ceranae*	Hypothetical protein (AAJ76_2600030604)	7.00E-38	*Nosema ceranae*	Hypothetical protein AAJ76_2600030604	2.00E-05
34	SSH-Forward-lib-53	668	*Nosema ceranae*	Forkhead hnf3 transcription factor (AAJ76_500091146)	0.00E+00	*Nosema ceranae*	Forkhead hnf3 transcription factor	7.00E-126
35	SSH-Forward-lib-54	305	*Nosema ceranae*	Histone H4 (AAJ76_150002436)	8.00E-127	*Nosema ceranae*	Histone H4	1.00E-39
36	SSH-Forward-lib-55	731	*Apis mellifera*	PREDICTED: *Apis mellifera* facilitated trehalose transporter Tret1-like (LOC724874)	0.00E+00	*Apis mellifera*	PREDICTED: facilitated trehalose transporter Tret1-like isoform X4	2.00E-07
37	SSH-Forward-lib-56	279	*Apis mellifera*	*Apis mellifera* heat shock protein cognate 4 (Hsc70-4) (NM_001160050)	2.00E-117	*Bombus terrestris*	Heat shock protein cognate 70	6.00E-50
38	SSH-Forward-lib-57	420	*Apis mellifera*	*Apis mellifera* heat shock protein cognate 4 (Hsc70-4) (NM_001160050)	7.00E-164	*Melipona quadrifasciata*	Heat shock 70 kDa protein cognate 4	2.00E-69
39	SSH-Forward-lib-59	231	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	4.00E-93	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	6.00E-35
40	SSH-Forward-lib-60	668	*Nosema ceranae*	Forkhead hnf3 transcription factor (AAJ76_500091146)	0.00E+00	*Nosema ceranae*	Forkhead hnf3 transcription factor	1.00E-124
41	SSH-Forward-lib-61	811	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	1.00E-84
42	SSH-Forward-lib-62	163	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	1.00E-47	*Nosema ceranae*	Actin	6.00E-16

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

43	SSH-Forward-lib-63	651	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	4.00E-141
44	SSH-Forward-lib-65	534	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	0.00E+00	*Nosema ceranae*	Hypothetical protein (AAJ76_2000141845)	3.00E-87
45	SSH-Forward-lib-66	615	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	0.00E+00	*Nosema ceranae*	Actin	2.00E-133
46	SSH-Forward-lib-68	671	*Apis mellifera*	PREDICTED: *Apis mellifera* ankyrin repeat domain-containing protein 49-like (LOC408773)	0.00E+00	*Apis dorsata*	PREDICTED: ankyrin repeat domain-containing protein 49-like	3.00E-90
47	SSH-Forward-lib-70	163	*Nosema ceranae*	Hypothetical protein (AAJ76_500092411)	3.00E-49	*Nosema ceranae*	Actin	2.00E-16
48	SSH-Forward-lib-71	550	*Nosema ceranae*	Histone H3 (AAJ76_800037613)	0.00E+00	*Nosema ceranae*	Histone H3	4.00E-89
49	SSH-Forward-lib-73	248	*Apis mellifera*	PREDICTED: *Apis mellifera* ribosomal protein S7 (RpS7) (XM_624940) (LOC552564)	1.00E-101	*Bombus impatiens*	PREDICTED: 40S ribosomal protein S7	8.00E-22
50	SSH-Forward-lib-74	892	*Nosema ceranae*	14-3-3 protein 1-like protein (AAJ76_300017093)	0.00E+00	*Nosema ceranae*	14-3-3 protein 1-like protein	0.00E+00
51	SSH-Forward-lib-75	522	*Nosema ceranae*	Nucleolar transformer 2-like protein (AAJ76_1900029507)	0.00E+00	*Nosema ceranae*	Nucleolar transformer 2-like protein	2.00E-92
52	SSH-Forward-lib-76	521	*Apis mellifera*	PREDICTED: *Apis mellifera* uncharacterized (ncRNA) (LOC107964104)	0	-	-	-
53	SSH-Forward-lib-78	616	*Nosema ceranae*	Hypothetical protein (AAJ76_1900028347)	0.00E+00	*Nosema ceranae*	Hypothetical protein AAJ76_1900028347	7.00E-79
54	SSH-Forward-lib-79	293	*Nosema ceranae*	Hypothetical protein (AAJ76_1900028347)	5.00E-125	*Nosema ceranae*	Hypothetical protein AAJ76_1900028347	6.00E-51
55	SSH-Forward-lib-80	294	*Nosema ceranae*	Hypothetical protein (AAJ76_1900028347)	2.00E-124	*Nosema ceranae*	Hypothetical protein AAJ76_1900028347	6.00E-48
56	SSH-Forward-lib-82	226	*Nosema ceranae*	Adp-ribosylation factor (AAJ76_1600043340)	3.00E-100	*Nosema ceranae*	Adp-ribosylation factor	2.00E-40
57	SSH-Forward-lib-83	447	*Nosema ceranae*	DNA-directed rna polymerase ii 16kda polypeptide (AAJ76_3400016364)	0.00E+00	*Nosema ceranae*	DNA-directed rna polymerase ii 16kda polypeptide	6.00E-78
58	SSH-Forward-lib-84	227	*Nosema ceranae*	Histone H3 (AAJ76_800037613)	2.00E-90	*Nosema ceranae*	Histone H3	2.00E-38

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

59	SSH-Forward-lib-86	346	*Nosema ceranae*	Sec61beta (AAJ76_5600014344)	7.00E-129	*Nosema ceranae*	Sec61beta	2.00E-37
60	SSH-Forward-lib-87	354	*Nosema ceranae*	Heat shock protein 90 (AAJ76_1170002375)	2.00E-151	*Nosema ceranae*	Heat shock protein 90	2.00E-60
61	SSH-Forward-lib-88	886	*Nosema ceranae*	Elongation factor 1 (AAJ76_1200007387)	0.00E+00	*Nosema ceranae*	Elongation factor 1	0.00E+00
62	SSH-Forward-lib-90	252	*Nosema ceranae*	Ubiquitin (AAJ76_1300017036)	3.00E-90	*Nosema ceranae*	Ubiquitin	3.00E-38
63	SSH-Forward-lib-94	529	*Nosema ceranae*	60s ribosomal protein L23 (AAJ76_200068052)	0.00E+00	*Nosema ceranae*	60S ribosomal protein L23	1.00E-101
64	SSH-Forward-lib-95	320	*Hottentotta judaicus*	Hottentotta judaicus clone Hj0016 gamma(m)-buthitoxin-Hj1c pseudogene mRNA (HQ288097)	1.00E-14	*Hottentotta judaicus*	Gamma-buthitoxin-Hj1a	3.00E-18
65	SSH-Forward-lib-96	401	*Nosema ceranae*	Histone H4 (AAJ76_150002436)	1.00E-151	*Nosema ceranae*	Histone h4	5.00E-52

2

3 **Table 3.** Analysis of 53 differential expressed sequence tags (ESTs) identified from reverse cDNA library

No	EST number	Insert Sequence length (bp)	NCBI_BLAST_N			NCBI_BLAST_X		
			Species	Gene name	E-value	Species	Gene name	E-value
1	SSH-Reverse-lib-1	387	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1, partial sequence (GU132310)	3.00E-03	*Elephantulus edwardii*	histone deacetylase complex subunit SAP130-like	1.5
2	SSH-Reverse-lib-2	216	*Apis mellifera*	PREDICTED: Apis mellifera facilitated trehalose transporter Tret1-like (LOC724874)	3.00E-50	*Escherichia coli*	hypothetical protein ECP02989421_5334	7.00E-90
3	SSH-Reverse-lib-3	156	*Apis mellifera*	PREDICTED: Apis mellifera chitin synthase chs-2 (LOC412215)	6.00E-51	-	-	-
4	SSH-Reverse-lib-4	467	*Apis mellifera*	PREDICTED: Apis mellifera histone H2A-like (LOC552322)	0.00E+00	*Nasonia vitripennis*	PREDICTED: histone H2A	4.00E-57
5	SSH-Reverse-lib-5	829	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC724199 (LOC724199)	0.00E+00	*Apis dorsata*	PREDICTED: extensin-like	2.00E-28
6	SSH-Reverse-lib-6	331	*Bombus impatiens*	PREDICTED: Bombus impatiens zinc finger protein 665-like (LOC100740332)	1.00E-39	-	-	-

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

7	SSH-Reverse-lib-7	488	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC411774 (LOC411774)	0.00E+00	*Apis mellifera*	PREDICTED: protein SMG9-like	3.00E-95
8	SSH-Reverse-lib-10	382	*Apis mellifera syriaca*	*Apis mellifera* syriaca mitochondrion (KP163643)	5.00E-38	-	-	-
9	SSH-Reverse-lib-12	233	*Megachile rotundata*	PREDICTED: *Megachile rotundata* probable maleylacetoacetate isomerase 2 (LOC100882531)	2.00E-14	-	-	-
10	SSH-Reverse-lib-14	192	*Apis mellifera*	PREDICTED: Apis mellifera caspase-like (LOC411381)	3.00E-69	*Apis mellifera*	PREDICTED: caspase-like	7.00E-29
11	SSH-Reverse-lib-15	546	*Apis dorsata*	PREDICTED: Apis dorsata nuclease-sensitive element-binding protein 1-like (LOC102681001)	0.00E+00	-	-	-
12	SSH-Reverse-lib-16	218	*Ceratitis capitata*	Ceratitis capitata clone 17a mRNA (DQ406807)	0.019	-	-	-
13	SSH-Reverse-lib-17	629	*Apis florea*	PREDICTED: Apis florea histone H3.3 (LOC100869036)	4.00E-30	-	-	-
14	SSH-Reverse-lib-18	166	*Hottentotta judaicus*	Hottentotta judaicus clone Hj0016 gamma(m)-buthitoxin-Hj1c pseudogene mRNA (HQ288097)	2.00E-12	-	-	-
15	SSH-Reverse-lib-19	246	*Apis mellifera*	PREDICTED: Apis mellifera probable alpha-ketoglutarate-dependent dioxygenase ABH4-like (LOC551104)	6.00E-98	*Apis mellifera*	PREDICTED: alpha-ketoglutarate-dependent dioxygenase alkB homolog 4-like	1.00E-25
16	SSH-Reverse-lib-20	658	*Apis mellifera*	PREDICTED: Apis mellifera facilitated trehalose transporter Tret1-like (LOC724874)	0.00E+00	*Apis mellifera*	PREDICTED: facilitated trehalose transporter Tret1-like isoform X1	7.00E-81
17	SSH-Reverse-lib-21	976	*Boechera divaricarpa*	Boechera divaricarpa GSS (HF949778)	0.00E+00	*Apis mellifera*	PREDICTED: transcription initiation factor IIA subunit 1 isoform 1	8.00E-48

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

18	SSH-Reverse-lib-22	469	*Apis mellifera*	PREDICTED: Apis mellifera ribosomal protein L3 (RpL3) (LOC552445)	0.00E+00	*Apis mellifera*	PREDICTED: 60S ribosomal protein L3	1.00E-80
19	SSH-Reverse-lib-23	276	*Aeromonas hydrophila*	Aeromonas hydrophila class I integron aminoglycoside adenylyltransferase (aadA4a) gene (DQ536502)	2.00E-16	*Escherichia coli*	hypothetical protein ECDG_04665	3.00E-29
20	SSH-Reverse-lib-25	482	*Apis mellifera*	PREDICTED: Apis mellifera UNC93-like protein MFSD11-like (LOC552407)	0.00E+00	*Apis mellifera*	PREDICTED: UNC93-like protein MFSD11-like	8.00E-77
21	SSH-Reverse-lib-26	273	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC726855 (LOC726855)	1.00E-108	*Apis mellifera*	PREDICTED: vegetative cell wall protein gp1-like isoform X1	8.00E-30
22	SSH-Reverse-lib-28	440	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC102653963 (LOC102653963)	0	-	-	-
23	SSH-Reverse-lib-32	556	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC724386 (LOC724386)	0.00E+00	*Apis mellifera*	PREDICTED: protein NPC2 homolog	1.00E-78
24	SSH-Reverse-lib-35	269	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC726855 (LOC726855)	1.00E-100	*Apis mellifera*	PREDICTED: vegetative cell wall protein gp1-like isoform X1	3.00E-25
25	SSH-Reverse-lib-37	701	*Pseudomonas simiae*	Pseudomonas simiae strain WCS417 genome (CP007637)	0.83	-	-	-
26	SSH-Reverse-lib-38	318	*Apis mellifera*	PREDICTED: Apis mellifera RNA-dependent helicase p72 (LOC411250)	6.00E-131	-	-	-
27	SSH-Reverse-lib-39	233	*Megachile rotundata*	PREDICTED: Megachile rotundata probable maleylacetoacetate isomerase 2 (LOC100882531)	2.00E-14	-	-	-
28	SSH-Reverse-lib-41	454	*Apis mellifera*	PREDICTED: Apis mellifera chitotriosidase-1-like (LOC100577156)	3.00E-30	*Apis mellifera*	PREDICTED: chitotriosidase-1-like isoform X3	2.00E-10
29	SSH-Reverse-lib-45	377	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	2.00E-04	*Elephantulus edwardii*	PREDICTED: histone deacetylase complex subunit SAP130-like	0.99

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

30	SSH-Reverse-lib-49	386	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	2.00E-04	*Elephantulus edwardii*	PREDICTED: WD repeat-containing protein 27	4.4
31	SSH-Reverse-lib-52	247	*Apis mellifera*	PREDICTED: Apis mellifera RNA-dependent helicase p72 (LOC411250)	1.00E-101	*Apis dorsata*	DEAD-box ATP-dependent RNA helicase 20-like	0
32	SSH-Reverse-lib-53	58	*Aeromonas hydrophila*	Aeromonas hydrophila class I integron aminoglycoside adenylyltransferase (aadA4a) gene	2.00E-12	*Escherichia coli*	hypothetical protein ECP02989421_5334	5.00E-92
33	SSH-Reverse-lib-58	213	*Apis mellifera*	PREDICTED: Apis mellifera phosphotidylinositol 3 kinase 21B ortholog (LOC408577)	4.00E-80	*Curtobacterium sp.*	hypothetical protein	4.8
34	SSH-Reverse-lib-59	327	*Apis mellifera*	PREDICTED: Apis mellifera ribosomal protein S7 (RpS7) (LOC552564)	2.00E-138	*Apis mellifera*	PREDICTED: 40S ribosomal protein S7	2.00E-58
35	SSH-Reverse-lib-60	199	*Apis mellifera*	PREDICTED: Apis mellifera solute carrier organic anion transporter family member 5A1-like (LOC409650)	1.00E-60	-	-	-
36	SSH-Reverse-lib-61	237	*Apis mellifera*	PREDICTED: Apis mellifera Jun-related antigen (Jra) (LOC726289)	7.00E-91	*Salmonella enterica*	hypothetical protein, partial	2.4
37	SSH-Reverse-lib-63	274	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC726855 (LOC726855)	1.00E-107	*Apis mellifera*	PREDICTED: vegetative cell wall protein gp1-like isoform X1	3.00E-18
38	SSH-Reverse-lib-67	377	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	0.003	*Danio rerio*	glutamate receptor 3 precursor	5.3
39	SSH-Reverse-lib-68	386	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	2.00E-04	-	-	-
40	SSH-Reverse-lib-69	376	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	0.003	*Nothobranchius furzeri*	PREDICTED: histone deacetylase complex subunit SAP130	4.7
41	SSH-Reverse-lib-70	275	*Apis mellifera*	PREDICTED: Apis mellifera upstream activation factor subunit spp27-like (LOC408520)	3.00E-109	-	-	-

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

42	SSH-Reverse-lib-71	377	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	2.00E-04	*Elephantulus edwardii*	PREDICTED: WD repeat-containing protein 27	4.9
43	SSH-Reverse-lib-72	329	*Apis mellifera*	PREDICTED: Apis mellifera transcription factor mblk-1-like (Mblk-1)	1.00E-139	-	-	-
44	SSH-Reverse-lib-74	563	*Apis dorsata*	PREDICTED: Apis dorsata uncharacterized (LOC102671476)	0.00E+00	*Apis dorsata*	PREDICTED: uncharacterized protein LOC102671476	6.00E-95
45	SSH-Reverse-lib-75	281	*Apis mellifera*	Phosphatidylinositol 4-phosphate 5-kinase type-1 alpha (LOC724991)	4.00E-115	-	-	-
46	SSH-Reverse-lib-79	946	*Apis mellifera*	PREDICTED: Apis mellifera protein FAM76A-like (LOC408294)	0.00E+00	*Apis mellifera*	PREDICTED: ATPase family AAA domain-containing protein 2-like	1.00E-63
47	SSH-Reverse-lib-84	383	*Apis mellifera*	PREDICTED: Apis mellifera ribosomal protein L21 (RpL21) (LOC726056)	3.00E-134	*Apis mellifera*	PREDICTED: 60S ribosomal protein L21	3.00E-47
48	SSH-Reverse-lib-85	142	*Uncultured prokaryote*	Uncultured prokaryote clone GFUGF50TF genomic sequence (EU923588)	0.003	-	-	-
49	SSH-Reverse-lib-86	55	*Pseudomonas aeruginosa*	integron metallo-beta-lactamase VIM-2 (blaVIM-2)	8.00E-11	-	-	-
50	SSH-Reverse-lib-87	404	*Apis mellifera*	PREDICTED: Apis mellifera uncharacterized LOC412243 (LOC412243)	3.00E-167	-	-	-
51	SSH-Reverse-lib-92	532	*Hottentotta judaicus*	gamma(m)-buthitoxin-Hj1c pseudogene mRNA (HQ288097)	2.00E-14	*Apis cerana cerana*	hypothetical protein APICC_07509	5.6
52	SSH-Reverse-lib-94	487	*Colletes stepheni*	Colletes stepheni isolate CSTM4 internal transcribed spacer 1 (GU132310)	3.00E-04	*Nothobranchius furzeri*	PREDICTED: histone deacetylase complex subunit SAP130	4.2
53	SSH-Reverse-lib-96	55	*Pseudomonas aeruginosa*	integron metallo-beta-lactamase VIM-2 (blaVIM-2	2.00E-10	-	-	-

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

Supplementary Table 1. Primer sets list for this study

No.	Gene	Primer Name	Sequence (5' to 3')	Tm (℃)	Note
1	SSU rDNA	18f	CACCAGGTTGATTCTGCC	50	Baker et al., 1995
		1537r	TTATGATCCTGCTAATGGTTC		
2	18S	143F	TGCCTTATCAGCTNTCGATTGTAG	50	Lo et al., 1997
		145R	TTCAGNTTTGCAACCATACTT CCC		
3	NCER_100435	F	GCACACCTGGATCAACAGAGAA	59	This study
		R	CTTAAGAACAAGTGCACCAGATCTA		
4	NCER_101664	F	GCTCTTGTTGTTGACAATGGTTC	59	This study
		R	AGCATCATCCCCAGCAAATC		
5	NCER_100566	F	TTGATTCTAGCCACTTTTACCGA	59	This study
		R	ACCCGAGAAGCTGCAGATATTAA		
6	NCER_100064	F	CGTGCAAATTAAAAGAAGATATTATTAA	59	This study
		R	CCATCATACCTGCAAAAGTCATAG		
7	NCER_101665	F	CAATACACGCAAAGAAAACCAAAT	60	This study
		R	TCCTTCAGCGGATTCTAAAGC		
8	NCER_101348	F	ATGGAAAACGATGAATACGCTG	60	This study
		R	GCCATGTCTTCATATCTTTCTGCTA		
9	NCER_100249	F	GTATACCACCAGATCAACAGAGACTT	59	This study
		R	TCATTAAGTGTTCTTCCATCTTCCA		
10	NCER_101591	F	CTACAATACAATCACTTCCAGGTCTT	59	This study
		R	GTGCGACTCCCAATAAACCAT		
11	NCER_101194	F	TGGATGAACTTTGTTGTTGGTGT	59	This study
		R	TTATGTCCTTGGATTAGTTCTCTTGA		
12	18S-Qpcr	Hb-18s-F	GTAACCCGTTGAACCCCATT	59	Schmittgen et al., 2000
		Hb-18s-R	CCATCCAATCGGTAGTAGCG		

Supplementary Table 2. Gene ontology analysis of forward library and reverse library

GO Classification	GO term	No. of EST	
		Forward library	Reverse library
Molecular Function	catalytic activity (GO:0003824)	5	4
	binding (GO:0005488)	3	4
	transporter activity (GO:0005215)	3	3
	molecular function regulator (GO:0098772)	1	1
	structural molecule activity (GO:0005198)	-	1
	transcription regulator activity (GO:0140110)	-	1
Biological Process	cellular process (GO:0009987)	4	4
	response to stimulus (GO:0050896)	2	2
	cellular component organization or biogenesis (GO:0071840)	1	1
	metabolic process (GO:0008152)	-	5

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

	biological regulation (GO:0065007)	-	2
	localization (GO:0051179)	-	1
Cellular Component	membrane (GO:0016020)	4	4
	cell (GO:0005623)	3	5
	protein-containing complex (GO:0032991)	2	1
	organelle (GO:0043226)	1	3

4. Conclusions

In this study, 28 honey bee genes and 24 genes *N. ceranae* were identified as DEGs after honey bees infected with *N. ceranae* via SSH approach. For the host DEGs, a high percentage of DEGs involved in catalytic activity and metabolic processes revealed the host gene expression change after *N. ceranae* infection might lead to the unbalance of physiological mechanism. Of 24 *N. ceranae* genes, nine DEGs were subjected to real-time quantitative reverse transcription PCR (real-time qRT-PCR) for validation and all of them showed high expression levels during different time post infections. In terms the results of validation, these genes revealed highly potential for the detection of different stages of nosemosis in the field. It should be also noted that one *N. ceranae* gene (AAJ76_1600052943, sr22) showed the highest expression level after infection. This gene might be further developed as a biomarker for early nosemosis detection. The data from this study could provide information on the pathological effects of *N. ceranae* and new insight for further applications on honey bee pathogen detections.

5. Patent

Taiwan, I481618. Lo, C. F., Wang, C. H. and Nai, Y. S. *Molecular Markers for nosema infection.*

Author Contributions: Conceptualization, Y.W. Chen; Y.S. Nai; Funding acquisition, Y.W. Chen; Y.S. Nai; Data curation, Z.T. Chang; C.Y. Ko; M.R. Yen; Formal analysis, M.R. Yen; Investigation, Z.T. Chang; Y.S. Nai; Methodology, C.Y. Ko; Y.S. Nai; Supervision, Y.S. Nai; Validation, Z.T. Chang; Writing – original draft, Z.T. Chang; Y.S. Nai; Writing – review & editing, M.R. Yen; Y.W. Chen; Y.S. Nai;

Funding: This work is supported by the Bureau of Animal and Plant Health Inspection and Quarantine, the Council of Agriculture and the Grant MOST 107-2313-B-197 -004 -MY3 from the Ministry of Science and Technology, Taiwan.

Acknowledgments: In memory of Prof. Chung-Hsiung Wang, who was a leadership of this article. We also thanks Ms. Po-Ya Hsu, Department of Biotechnology and Animal Science, National Ilan University, Taiwan for kindly help on the sequence analyses and technical support. This article

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Johnson, R.M., et al., *Changes in transcript abundance relating to colony collapse disorder in honey bees (Apis mellifera).* Proc Natl Acad Sci U S A, 2009. **106**(35): p. 14790-5.
2. Sanchez-Bayo, F. and K. Goka, *Pesticide residues and bees--a risk assessment.* PLoS One, 2014. **9**(4): p. e94482.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

3. Sciences, N.A.o., ed. *Status of Pollinators in North America*. National Academies Press, Washington, DC. ed. 2007.
4. Wenning, R.J., *The status of pollinators.* Integr Environ Assess Manag, 2007. **3**(3): p. 309.
5. Cox-Foster, D.L., et al., *A metagenomic survey of microbes in honey bee colony collapse disorder.* Science, 2007. **318**(5848): p. 283-7.
6. vanEngelsdorp, D., et al., *An estimate of managed colony losses in the winter of 2006-2007: A report commissioned by the apiary inspectors of America.* Am. Bee J., 2007. **147**: p. 599-603.
7. Higes, M., et al., *Honeybee colony collapse due to Nosema ceranae in professional apiaries.* Environ Microbiol Rep, 2009. **1**(2): p. 110-3.
8. Tozkar, C.O., et al., *Metatranscriptomic analyses of honey bee colonies.* Front Genet, 2015. **6**: p. 100.
9. Higes, M., et al., *The differential development of microsporidia infecting worker honey bee (Apis mellifera) at increasing incubation temperature.* Environ Microbiol Rep, 2010. **2**(6): p. 745-8.
10. Adl, S.M., et al., *The new higher level classification of eukaryotes with emphasis on the taxonomy of protists.* J Eukaryot Microbiol, 2005. **52**(5): p. 399-451.
11. Higes, M., et al., *Experimental infection of Apis mellifera honeybees with Nosema ceranae (Microsporidia).* J Invertebr Pathol, 2007. **94**(3): p. 211-7.
12. Fries, I., et al., *Nosema ceranae n. sp. (Microspora, Nosematidae), morphological and molecular characterization of a microsporidian parasite of the Asian honey bee Apis cerana (Hymenoptera, Apidae).* Eur J Protistol., 1996. **32**: p. 356-365.
13. Furgala, B.F. and E.C. Mussen, *Honey Bee Pests, Predators and Diseases*. Protozoa, ed. R.A. Morse and R. Nowogrodzki. 1990, New York, NY, USA: Cornell University Press. 49-63.
14. Huang, W.F., et al., *A Nosema ceranae isolate from the honeybee Apis mellifera.* Apidologie, 2007. **38**: p. 30-37.
15. Higes, M., R. Martin, and A. Meana, *Nosema ceranae, a new microsporidian parasite in honeybees in Europe.* J Invertebr Pathol, 2006. **92**(2): p. 93-5.
16. Martin-Hernandez, R., et al., *Nosema ceranae in Apis mellifera: a 12 years postdetection perspective.* Environ Microbiol, 2018. **20**(4): p. 1302-1329.
17. Martin-Hernandez, R., et al., *Outcome of colonization of Apis mellifera by Nosema ceranae.* Appl Environ Microbiol, 2007. **73**(20): p. 6331-8.
18. Higes, M., et al., *A preliminary study of the epidemiological factors related to honey bee colony loss in Spain.* Environ Microbiol Rep, 2010. **2**(2): p. 243-50.
19. Martin-Hernandez, R., et al., *Comparison of the energetic stress associated with experimental Nosema ceranae and Nosema apis infection of honeybees (Apis mellifera).* Parasitol Res, 2011. **109**(3): p. 605-12.
20. Botias, C., et al., *Nosema spp. infection and its negative effects on honey bees (Apis mellifera iberiensis) at the colony level.* Vet Res, 2013. **44**: p. 25.
21. Dussaubat, C., et al., *Gut pathology and responses to the microsporidium Nosema ceranae in the honey bee Apis mellifera.* PLoS One, 2012. **7**(5): p. e37017.
22. Szumowski, S.C. and E.R. Troemel, *Microsporidia-host interactions.* Curr Opin Microbiol, 2015. **26**: p. 10-6.
23. Diatchenko, L., et al., *Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries.* Proc Natl Acad Sci U S A, 1996. **93**(12): p. 6025-30.

Peer-reviewed version available at Insects 2020, 11, 199; doi:10.3390/insects11030199

24. Tsai, S.J., et al., *The characterization of microsporidian isolates (Nosematidae: Nosema) from five important lepidopteran pests in Taiwan.* J Invertebr Pathol, 2003. **83**(1): p. 51-9.
25. Wang, C.Y., et al., *An Endoreticulatus species from Ocinara lida (Lepidoptera: Bombycidae) in Taiwan.* J Invertebr Pathol, 2005. **89**(2): p. 123-35.
26. Malone, L.A., H.A. Giacon, and M.R. Newton, *Comparison of the responses of some New Zealand and Australian honey bees (Apis mellifera L) to Nosema apis Z.* Apidologie, 1999. **26**: p. 495-502.
27. Tsai, S.J., et al., *Complete sequence and structure of ribosomal RNA gene of Heterosporis anguillarum.* Dis Aquat Organ, 2002. **49**(3): p. 199-206.
28. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
29. Aufauvre, J., et al., *Transcriptome analyses of the honeybee response to Nosema ceranae and insecticides.* PLoS One, 2014. **9**(3): p. e91686.
30. Antunez, K., et al., *Immune suppression in the honey bee (Apis mellifera) following infection by Nosema ceranae (Microsporidia).* Environ Microbiol, 2009. **11**(9): p. 2284-90.
31. Chaimanee, V., et al., *Differential expression of immune genes of adult honey bee (Apis mellifera) after inoculated by Nosema ceranae.* J Insect Physiol, 2012. **58**(8): p. 1090-5.
32. Li, Z., et al., *Molecular and biochemical responses in the midgut of the silkworm, Bombyx mori, infected with Nosema bombycis.* Parasit Vectors, 2018. **11**(1): p. 147.
33. Hyland, E.M., et al., *Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae.* Mol Cell Biol, 2005. **25**(22): p. 10060-70.
34. Porter, G.W., F.R. Khuri, and H. Fu, *Dynamic 14-3-3/client protein interactions integrate survival and apoptotic pathways.* Semin Cancer Biol, 2006. **16**(3): p. 193-202.
35. Zuehlke, A.D., M.A. Moses, and L. Neckers, *Heat shock protein 90: its inhibition and function.* Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1738).
36. Weigel, D. and H. Jäckle, *The fork head domain: A novel DNA binding motif of eukaryotic transcription factors?* Cell, 1990. **63**(3): p. 455-456.
37. Partridge, L. and J.C. Bruning, *Forkhead transcription factors and ageing.* Oncogene, 2008. **27**(16): p. 2351-63.
38. Qin, X., et al., *Genome sequences of the honey bee pathogens Paenibacillus larvae and Ascosphaera apis.* Insect Mol Biol, 2006. **15**(5): p. 715-8.
39. Cornman, R.S., et al., *Genomic analyses of the microsporidian Nosema ceranae, an emergent pathogen of honey bees.* PLoS Pathog, 2009. **5**(6): p. e1000466.