A multi-generation risk assessment of Cry1F on the non-target soil organism *Folsomia candida* (Collembola) based on whole transcriptome profiling (#32777)

First submission

Editor guidance

Please submit by **13 Dec 2018** for the benefit of the authors (and your $200 publishing discount).

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

5 Figure file(s)
3 Table file(s)
2 Raw data file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- Data is robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

A multi-generation risk assessment of Cry1F on the non-target soil organism *Folsomia candida* (Collembola) based on whole transcriptome profiling

Cheng-Wang Huang [1], **Wan-Jun Chen** [1], **Xin Ke** [1], **Yun-he Li** Corresp., [2], **Yun-Xia Luan** Corresp. [1, 3]

[1] Key Laboratory of Insect Developmental and Evolutionary Biology, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China
[2] State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China
[3] Institute of Insect Science and Technology & School of Life Sciences, South China Normal University, Guangzhou, China

Corresponding Authors: Yun-he Li, Yun-Xia Luan
Email address: liyunhe@caas.cn, yxluan@scnu.ac.cn

The *Bacillus thuringiensis* (*Bt*) toxin Cry1F has been used to develop insect-resistant genetically engineered (IRGE) crops, but its potential risk to the non-target soil invertebrates has not been evaluated well. Here, we conducted a laboratory-based multi-generation risk assessment of Cry1F for collembolan *Folsomia candida*, an important representative of soil arthropods, in terms of survival, reproduction, and differentially expressed genes (DEGs) identified from whole transcriptome profiles. Our results demonstrated that Cry1F was continuously ingested by collembolans over three consecutive generations, but it did not affect the survival or reproduction of *F. candida*. There were no significant differences in the global gene expression between *F. candida* fed diets with and without Cry1F, and no consistent co-expressed DEGs over three generations. In addition, Cry1F did not obviously alter the expression profiles of seven sensitive biological markers. Our composite data indicate that Cry1F had no long-term harmful effect on collembolan *F. candida*.

A multi-generation risk assessment of Cry1F on the non-target soil organism *Folsomia candida* (Collembola) based on whole transcriptome profiling

Cheng-Wang Huang[1], Wan-Jun Chen[1], Xin Ke[1], Yun-he Li[2], Yun-Xia Luan[1, 3]

[1]Key Laboratory of Insect Developmental and Evolutionary Biology, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

[2]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

[3] Guangdong Key Laboratory of Insect Development Science and Technology, Institute of Insect Science and Technology & School of Life Sciences, South China Normal University, Guangzhou, China.

Corresponding Authors:

Yun-Xia Luan[1,3] and Yun-he Li[2]

Email address: yxluan@scnu.ac.cn; liyunhe@caas.cn

ABSTRACT

The *Bacillus thuringiensis* (*Bt*) toxin Cry1F has been used to develop insect-resistant genetically engineered (IRGE) crops, but its potential risk to the non-target soil invertebrates has not been evaluated well. Here, we conducted a laboratory-based multi-generation risk assessment of Cry1F for collembolan *Folsomia candida*, an important representative of soil arthropods, in terms of survival, reproduction, and differentially expressed genes (DEGs) identified from whole transcriptome profiles. Our results demonstrated that Cry1F was continuously ingested by collembolans over three consecutive generations, but it did not affect the survival or reproduction of *F. candida*. There were no significant differences in the global gene expression between *F. candida* fed diets with and without Cry1F, and no consistent co-expressed DEGs over three generations. In addition, Cry1F did not obviously alter the expression profiles of seven sensitive biological markers. Our composite data indicate that Cry1F had no long-term harmful effect on collembolan *F. candida*.

Keywords: Environmental risk assessment, Cry1F, Collembolan, Consecutive generations, Transcriptome analysis.

INTRODUCTION

Over the past 20 years, genetically engineered (GE) crops have been widely planted throughout the world. The most common of these are insect-resistant genetically engineered (IRGE) crops, mainly including IRGE cotton, maize, tobacco and others (*ISAAA, 2016*). All the current commercialized IRGE crops express *cry* or *vip* genes derived from the bacterium *Bacillus thuringiensis* (*Bt*), and these genes encode insecticidal proteins targeting lepidopteran or coleopteran insect pests (*Liu et al., 2016*). Although the planting of IRGE crops will greatly reduce the use of broad-spectrum insecticides, the potential risks of IRGE crops to the environment and human health must be assessed before they are commercialized. One of the environmental risks associated with the planting of IRGE crops is their potential negative effect on non-target organisms (NTOs) (*Shelton, Zhao & Roush, 2002*), which follows a tiered approach from laboratory to field (*EFSA, 2010; Li et al., 2014; Romeis et al., 2008*). To date, many laboratory and field tests have been conducted with many NTOs belonging to a range of functional groups, including herbivores, predators, parasitoids, and detritivores (*Gao et al., 2018; Jia et al., 2016; Tian et al, 2014*).

Because of their large numbers in soil (10^4-10^5 m^{-2}) and substantial contribution to the decomposition of plant residues, detritivore's collembolans are an important group of soil arthropods (*Hopkin, 1997*). Moreover, collembolans are quite sensitive to many soil pollutants, such as heavy metal ions, organic pollutants, and insecticides (*Buch et al., 2016; Chen et al., 2015; Zortéa et al., 2015*). Collembolans are exposed to *Bt* toxins when IRGE crop residues decompose (*Li et al., 2007; Valldor et al., 2015*), and negative effects of *Bt* toxins on

collembolans would be detrimental to the entire agro-ecosystem. The parthenogenetic collembolan *Folsomia candida* is a standard test organism that is often used for toxicological, biosafety, and environmental assessment. Most previous risk assessments didn't find any harmful effects of various *Bt* toxins on collembolans: *Sims & Martin (1997)* found CryIA(b), CryIA(c), CryIIA, and CryIIIA present in transgenic crop plants do not reduced survival or reproduction of collembolan *F. candida* and *Xenylla grisea*; *Yang et al. (2015, 2018)* reported that Cry1C and Cry2A toxins were not toxic to *F. candida*; and *Zhang et al. (2017a)* found that Cry1Ab/Cry2Aj-containing *Bt* maize pollen did not reduce *F. candida* fitness or induce any sublethal effects. However, these laboratory risk assessments only examined one generation of *F. candida*, i.e., they did not evaluate the potential risk to collembolan progeny.

The possibility that *Bt* plants might have long-term effects on the soil biota remains a major concern (*Icoz & Stotzky, 2008*), but such "chronic toxicity" has rarely been studied (*Clark, Phillips & Coats, 2005*). In multi-generation laboratory tests, *Bakonyi et al. (2011)* and *Szabó, Seres & Bakonyi (2017)* found that *F. candida* that were fed *Bt* maize leaves had some alterations in life-history traits and reproduction. To our knowledge, these are the only studies that evaluated the long-term potential risk of *Bt* toxins to collembolans. Additional studies are needed.

In assessing the effects of *Bt* toxins on collembolans, most previous studies have used physiological indices (survival rates, reproductive rates, developmental duration, etc.) or biological markers (enzyme activity, midgut bacterial diversity, etc.). These traditional methods of risk assessment may not detect slight changes at the molecular level. RNA sequencing (RNA-

Seq), which is useful for detecting genes that are differentially expressed in response to environmental stresses, can provide a whole transcriptome profile of NTOs to check stress-responsive changes in gene expression and to evaluate the potential risk of IRGE crops. By comparative transcriptome analysis, *Xu et al. (2017)* found that the expression of > 2,000 genes of larvae of the fish *Coryphaena hippurus* was significantly altered after exposure to Deepwater Horizon oil. Based on RNA-Seq data, *Zhang et al. (2017b)* annotated 32,268 unigenes of the Chinese green mussel (*Perna viridis*) and identified 9,048 differentially expressed genes (DEGs) between exposed and non-exposed groups to cadmium, suggesting a sensitive response of the mussel transcriptome to cadmium.

The *cry1F* has been transformed to maize, cotton, and tobacco, and previous studies have found no harmful effects of Cry1F on some tested non-target organisms (*Kim et al., 2012; Tian et al., 2014*). However, its risk to collembolans has not been well evaluated except for several field tests on the collembolan community (*Higgins et al., 2009; Marques et al., 2018*). In this study, we performed a dietary exposure experiment (DEE) with *F. candida* that was fed an artificial diet containing or not containing purified Cry1F protein for three consecutive generations. We compared the effects of these two diets on the survival, reproduction, and transcriptome changes of *F. candida* for each generation. To our knowledge, this is the first multi-generation risk assessment of Cry1F on *F. candida*, and also the first attempt to use RNA-Seq to evaluate the impact of a *Bt* toxin on the gene expression of *F. candida*.

MATERIALS AND METHODS

Test organism.

The Danish strain of the parthenogenetic *Folsomia candida* Willem, 1902, was originally obtained from Aarhus University, Denmark, and has been cultured in our laboratory for over 10 years. *F. candida* is usually fed granulated dried baker's yeast (Hebei Mauri Foods Co., Ltd., Zhangjiakou, China) and was reared in the Petri dishes (90 mm x 18 mm) containing a 5-7 mm layer of a solidified mixture of plaster of Paris and activated charcoal (9:1 wt/wt, dissolved in distilled water, with about 270 ml of water/500 g of mixed powder) (Fig. 1A-C). The baker's yeast was placed on the surface of the plaster and was renewed weekly to reduce the growth of other fungi. Distilled water was added to the base as needed such that free water was present in the plaster pores but did not form a film on the plaster surface. These dishes with *F. candida* were kept in an artificial climate chamber in total darkness (20 ± 1°C, 80% relative humidity).

Dietary exposure experiment (DEE).

The artificial diet containing Cry1F (BT diet) was prepared as follows: 2 mg of purified Cry1F protein (Envirologix Inc., Portland, Maine, USA) and 4 g of baker's yeast granules were dissolved in 10 ml of distilled water; the preparation was fully mixed and then transferred into a plastic container. The final concentration of Cry1F protein in the BT diet was 500 μg/g, much higher than EC_{50} values of Cry1F against lepidopteran pests *Chilo suppressalis* (6 μg/g) and larvae of *Bombyx mori* (136 ng/g) (*Jiao et al., 2016*). After 27 h of lyophilization, the BT diet was ground into powder and stored at -70°C. As the control, baker's yeast granules without Cry1F (CT diet) were prepared in the same way.

Based on the Organisation for Economic Cooperation and Development (OECD) test protocol (*OECD, 2016*), we designed a multi-generation DEE on *F. candida* (Fig. 1D). During the experiment, all *F. candida* were cultured in the same artificial climate chamber with the culture conditions mentioned above. They were synchronized to the same stage (newborn juveniles) at the start of each generation in the experiment.

First, we transferred 80-100 randomly selected *F. candida* adults (about 4-6 weeks old) to a new Petri dish. After 72 h of oviposition, all adults were removed. The eggs began to hatch 10 days later. After 24 h of hatching, we selected synchronized newborn juveniles. They were transferred to a new Petri dish and were fed with pure baker's yeast for 10 days. Next, all juveniles were separated into two sets: the BT set were fed BT diet, and the CT set were fed CT diet. Each set was done with individually- and group- reared collembolans.

In the first-generation DEE, 60 juveniles were transferred to 60 small plastic Petri dishes (55 mm x 14 mm, plaster height: 5 mm), respectively, as the individually-reared collembolans. 30 ones were fed BT diet, and 30 ones were fed CT diet. The rest juveniles were evenly distributed into six plastic Petri dishes (85 mm x 16 mm, plaster height: 5 mm), with 60 to 80 collembolans per dish. Three of them were fed BT diet and the other three were fed CT diet. All diets were renewed every 2 days to avoid Cry1F protein degradation and the growth of other fungi.

The first generation was assessed after 25 days of DEE. For the individually-reared collembolans, the numbers of adults and juveniles in each treatment were counted to calculate the survival and reproduction rates. For the collembolans fed in groups, about 20 individuals per replicate were randomly collected for ELISA (enzyme-linked immunosorbent assay, see section

136 *ELISA* below), and 25-35 individuals per replicate were collected for RNA-Seq (see section *RNA*
137 *extraction, sequencing and unigene annotation* below).
138 The remaining collembolans fed in groups of CT and BT sets were used for synchronization of
139 the second generation. To maintain the same experimental conditions, the CT and BT sets were
140 synchronized independently, and their newborn juveniles were fed CT diet or BT diet directly
141 after birth. After 10 days, the second-generation DEE started. It was done in the same manner as
142 the first-generation DEE except that there were 40 replicates per set for individually-reared
143 collembolans and two replicates per set for group-reared collembolans. After 25 days of DEE,
144 survival and reproduction were evaluated using the individually-reared collembolans, and ELISA
145 and RNA-Seq were performed suing the group-reared collembolans as described for the first
146 generation. The remaining collembolans that were fed in groups in the second generation were
147 used to start the third generation. The procedure for the third-generation DEE was identical to
148 that for the second except that there were three rather than two replicates per set for the group-
149 reared collembolans.
150
151 **Survival and reproduction.**
152 The individually reared collembolans were used to evaluate survival and reproduction of *F.*
153 *candida* at the end of each generation. Several replicates were excluded from analysis: 1)
154 disappeared adult collembolans (5 in BT treatment and 5 in CT treatment over three
155 generations), which are assumed to have escaped because no corpse left; 2) eggs didn't hatch due
156 to overgrown fungi (6 in BT treatment and 5 in CT treatment over three generations).

157 With the aid of a stereoscopic microscope (Nikon, SMZ-10), the number of living adults was
158 determined as a measure of survival, and the number of juveniles produced was determined as a
159 measure of reproduction. Due to the wide variation in the reproduction data (7-41 juveniles per
160 individually-reared collembolan), the data were log-transformed prior to the statistical analysis.
161 Student's *t*-test was used to analyze the differences between the two treatments in each
162 generation. In addition, the data for reproduction of all BT and CT sets over three generations
163 were subjected to a two-way ANOVA, followed by the LSD (least significant difference) test,
164 with diet treatment, generation and their interaction as fixed factors. All of these statistical
165 analyses were performed with IBM SPSS Statistics 24 (version R24.0.0.0). Differences were
166 considered significant at $p < 0.05$.
167
168 **ELISA.**
169 The concentrations of Cry1F in the diets and in *F. candida* fed in groups were measured by
170 ELISA for each replicate in each generation. The QuantiPlate Kit for Cry1F (Envirologix Inc.,
171 Portland, Maine, USA) was used to detected Cry1F in a 2- to 4-mg sample of the fresh diet in
172 each replicate, in a 2- to 4-mg sample of the diet in each replicate after 2 days of feeding, and in
173 20 individuals of *F. candida* (2-5 mg) per replicate. All samples were fully ground with an
174 electric grinding rod and then extracted with PBST extraction buffer (phosphate-buffered saline
175 with Tween-20, pH 7.4). ELISA was performed according to the manufacturer's instructions. A
176 two-way ANOVA was used to compare Cry1F concentrations of three BT sets over three
177 generations by using the software IBM SPSS Statistics 24 (version R24.0.0.0). Differences were

178 considered significant at $p < 0.05$.

179

180 **RNA extraction, sequencing and unigene annotation.**

181 The miRNeasy® Mini Kit (Qiagen Biotech, Germany) was used to extract the total RNA from
182 25-35 individuals per replicate of the collembolans that were fed in groups for each generation.
183 RNA concentration and integrity were evaluated with Aglient 2100 Bioanalyzer (Agilent, USA).

184 cDNA library construction, Illumina sequencing, and *de novo* assembly of RNA samples were
185 carried out by Hangzhou 1gene Technology Co., Ltd. All cDNA libraries were constructed using
186 the NEBNext®Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA), following the protocol
187 described by the manufacturer. The libraries were sequenced with 150-bp paired-end reads on an
188 Illumina Hiseq 4000 platform (San Diego, CA, USA). After sequencing, the raw reads (about 6
189 GB of data for each replicate) were filtered to remove adaptor sequences, duplication sequences,
190 and low-quality sequences. The clean reads were *de novo* assembled into unigenes by using
191 Trinity and SOAPdenovo-Trans. Six public databases were used to annotate unigenes with
192 BLASTx or BLASTn (E-value < 10^{-5}), including NCBI Nr and Nt databases
193 (http://www.ncbi.nlm.nih.gov/), SwissProt (http://www.expasy.ch/sprot/), KEGG
194 (http://www.genome.jp/kegg/), COG (http://www.ncbi.nlm.nih.gov/COG/), and GO
195 (http://www.geneontology.org/).

196

197 **Differential gene expression and biological marker expression.**

198 The expression levels of all unigenes were calculated by using the FPKM (fragments per kb

199 per million fragments) method: FPKM = $(10^6 C)/(NL/10^3)$. The data of three or two replicates for
200 each set in every generation were averaged, and whole transcriptome gene expression of all CT
201 and BT sets were compared. DEGs were determined by the condition of FC (fold change, |log2
202 ratio| of FPKM (BT vs. CT)) ≥ 1 and FDR (false discovery rate) ≤ 0.05.
203 We carefully checked the expression of seven biological markers whose changes in expression
204 level are often used to indicate that the test organism is experiencing stress. These genes encode
205 the following proteins: the antioxidant-related enzymes catalase (CAT) (*Niu et al., 2017; Yousef,*
206 *Abdelfattah & Augustyniak, 2017*) and superoxide dismutase (SOD) (*Yang et al., 2015; Yousef,*
207 *Abdelfattah & Augustyniak, 2017*); the detoxification-related enzymes glutathione S-transferase
208 (GST) (*Niu et al., 2017; Oliveira et al., 2015*), carboxylesterase (CES) (*Niu et al., 2017; Yang et*
209 *al., 2015*), and glutathione reductase (GR) (*Yang et al., 2015*); a metallothionein-like motif
210 containing protein (MTC) (*Nakamori & Kaneko, 2013*), which is unique in *F. candida* and which
211 is very sensitive to different heavy metal ions; and the heat shock protein 70 (HSP70) (*Liu et al.,*
212 *2010*), which is sensitive to insecticides, drought, and other environmental stresses. The
213 expression profiles of unigenes annotated to these biological markers were examined, and their
214 expression levels were compared among different samples.
215 The software R (version 3.4.0) was used for the heatmaps (R package 'pheatmap'
216 (https://cran.r-project.org/web/packages/pheatmap/index.html)) and the hierarchical clustering
217 ('dist' and 'hclust' orders, distance method as 'euclidean' and cluster method as 'complete') of
218 all DEGs and biological markers over three generations (expression data of DEGs were
219 transformed by using the common logarithm of the counts plus 1).

220

221 **RESULTS**

222 **Cry1F concentrations in the diet and *F. candida*.**

223 ELISA measures showed that no Cry1F toxin was detected in the pure yeast diet or in the *F.*
224 *candida* that were fed the CT diet. The average concentration (mean ± SE) of Cry1F was 415 ±
225 15 µg/g in the fresh BT diet and 358 ± 57 µg/g in the BT diet after 2 days of feeding, indicating
226 that Cry1F was continually present in the experiment. The average concentrations of Cry1F in
227 collembolans that were fed the BT diet were 3.5 ± 1.3, 2.7 ± 1.4, and 5.2 ± 1.0 µg/g at the end of
228 the first, second, and third generation, respectively; these concentrations were not significantly
229 different based on a two-way ANOVA ($p = 0.429$), and suggest that the Cry1F was continuously
230 ingested by collembolans.

231

232 **Survival and reproduction of *F. candida*.**

233 All collembolans in both CT and BT sets survived and reached the adult stage for all three
234 generation, which suggests Cry1F has no any negative effect on collembolan survival. Over the
235 three generations, the mean (± SE) number of juveniles produced per individually fed
236 collembolan was 22.57 ± 0.635 in the CT set and 20.70 ± 0.590 in the BT set. According to
237 Student's *t*-test for each generation, the number of juveniles produced per collembolan did not
238 significantly differ between CT and BT sets for any of the three generations ($p = 0.912$, 0.114,
239 0.071 for generations 1, 2 and 3, respectively) (Fig. 2). Two-way ANOVA (with diet treatment,

generation, and their interaction as fixed factors, Table 1) showed that the mean number of juveniles produced per collembolan was significantly affected by generation ($p = 0.001$) but not by diet treatment ($p = 0.068$) or the interaction between diet treatment and generation ($p = 0.410$).

Transcriptome sequencing and annotation.

A total of 316,693,647 raw reads for 16 samples were generated, and deposited in the NCBI Sequence Read Archive (SRP132745). After data filtering and *de novo* assembly, 284,174,422 clean reads were assembled into 93,976 unigenes. The total length of these unigenes was 155,829,628 bp. The mean length of the unigenes was 1,658 bp, and N50 length was 3,292 bp. The sequence data of each type of sample were unbiased, and enough reads were obtained to perform gene expression analyses. The quantity and quality of the RNA sequencing data are shown in Table 2.

Of the total number of unigenes, 55,390 (58.94%), 17,230 (18.33%), 47,734 (50.79%), 43,284 (46.06%), 31,016 (33.00%), and 18,603 (19.80%) were annotated in NCBI Nr, NCBI Nt, SwissProt, KEGG, COG, and GO databases, respectively. Overall, 57,758 (61.46%) unigenes were annotated to known protein/nucleotide sequences.

Differentially expressed genes (DEGs) over three generations.

463 DEGs (0.49% of all unigenes) between CT and BT sets were identified for all three generations (Fig. 3A-C), including 211, 19, and 244 DEGs for the first, second, and third

generation, respectively (Fig. 3D). There was no consistent tendency of either up-regulation or down-regulation of DEGs, and there was no consistent co-expression of DEGs over the three generations (Fig. 3E), and only 6 DEGs (5 up-regulated and 1 down-regulated) were identified in both the first and third generations. 5 contra-regulated DEGs were detected in the second and third generations. The functions of these 11 DEGs are unknown.

The heatmap analysis of DEGs over three generations showed similar expression patterns of the CT and BT sets in each generation (Fig. 4A), and the hierarchical clustering of all samples demonstrated CT and BT samples of the same generation were always clustered together, instead of the CT nor the BT sets of different generations (Fig. 4B).

Biological marker expression profile.

A total of 855 unigenes for the seven biomarkers (CAT, GST, SOD, GR, CES, MTC, and HSP70) were annotated, but only five unigenes were differentially expressed between CT and BT sets, i.e., were up- or down-regulated by more than 2-fold in the BT set (Table 3): one unigene of CAT (CL2174.Contig5_All) was up-regulated in the third generation; three unigenes of CES (Unigene7933_All, CL5034.Contig2_All, and CL6365.Contig1_All) encoding carboxylesterase type B were up- or down-regulated in the first or the third generations; one unigene of HSP70 (Unigene15285_All) was greatly down-regulated, i.e., there was a 5-fold decrease in expression, in the first generation. However, the expression of the five DEGs was significantly different in BT vs. CT sets in only a specific generation, and none of those changes are consistent in all three generations (Table 3). Furthermore, the gene expression hierarchical

clustering of the seven markers (Fig. 5) clearly showed that the CT and BT sets from the same generation always clustered together, which confirms that Cry1F treatment did not affect the functions related to the seven biological markers.

DISCUSSION

Multi-generation laboratory tests are useful for detecting the chronic toxicity of pollutants on collembolans (*Campiche et al., 2007; van Gestel et al., 2017*). Most previous assessments of *Bt* toxins or crops on collembolans involved the assessment of short-term exposure (one generation) rather than chronic toxicity (*Yang et al., 2015, 2018; Zhang et al., 2017a*). With the long-term cultivation of *Bt* crops, however, toxicity may be undetectable in the first generation of collembolans but may increase in subsequent generations because of continuous exposure to *Bt* toxins. Therefore, it is necessary to test the effects of *Bt* toxins on non-target organisms with multi-generation toxicological assessments. Two previous laboratory-based multi-generation risk assessments of *Bt* toxins on collembolans (*Bakonyi et al., 2011; Szabó, Seres & Bakonyi, 2017*) found that feeding on leaves of *Bt* maize (*MON 810*) affected egg production, growth rate, and food preference in *F. candida*. In addition, *Yuan et al.* (*2011*) found that feeding on *Bt* rice decrease collembolan catalase activity. However, these results are difficult to explain (except to note that *Bt* and non-*Bt* plant tissues may differ in properties other than the presence or absence of *Bt* toxin). The current study used a diet that was identical except for the presence or absence

of *Bt* toxin. That *F. candida* is parthenogenetic facilitated the synchronization of generations and the multi-generation assessment.

In addition to compare *F. candida* survival and reproduction, the current study also considered the effects of *Bt* toxin on gene expression. Microarray-based detection methods have been widely used to evaluate the molecular effect of pollutants on test organisms (*Baillon et al., 2016; Roelofs et al., 2009; Yuan et al., 2014*). Compared with microarray, RNA-Seq is more comprehensive and is better able to detect potential risk, since it does not rely on a pre-designed complement sequence detection probe; it enables the identification of genetic variants; and it can quantify and profile overall gene expression, including rare and novel transcripts (*Montgomery et al., 2010; Sultan et al., 2008; Wang, Gerstein & Snyder, 2009*). If the pollutants are harmful to *F. candida*, they should create a constant stress and result in co-expression of DEGs at different exposure times (*Qiao et al., 2015*), but this was not the case in the current study. The expression pattern of DEGs was more similar between CT and BT sets in the same generation than between the same diet treatment of different generations (Fig. 4), suggesting that the detected DEGs may be random events perhaps caused by uncontrolled conditions in the experiment rather than by Cry1F. Moreover, the statistical analysis showed that *F. candida* reproduction significantly differed among generations but not between the diet treatments (Table 1), which further confirmed that Cry1F had no constant or cumulative effect on *F. candida*. Although we synchronized *F. candida* to stage before every generation and although we strictly controlled the culture conditions, there were probably some uncontrolled differences in season, diet quality, or other factors between generations.

322 Because it is rapid, simple, and sensitive, the biological marker assay is commonly used for
323 the monitoring of environmental contaminants. *Yousef, Abdelfattah & Augustyniak (2017)* found
324 fertilizer industry pollutants greatly affected the activities of SOD and CAT in the grasshopper
325 *Aiolopus thalassinus*. By assessing GST responses, *Oliveira et al. (2015)* assayed the risk of *F.*
326 *candida* exposure to carbamazepine. *Nakamori & Kaneko (2013)* measured the effect of Cd
327 exposure on *F. candida* by assessing the gene expression of MTC, which is a biomarker related
328 to heavy metal detoxification. Most previous studies only determined the gene expression
329 differences of one or several biomarkers by RT-qPCR. In our study, RNA-Seq provided a whole
330 transcriptome profile, and we were therefore able to screen a large variety of important
331 biomarker genes in *F. candida*, including some novel response genes.
332 In summary, we established a laboratory-based multi-generation risk assessment of Cry1F for
333 *F. candida*, and found Cry1F did not affect the survival or reproduction of *F. candida* over three
334 consecutive generations, and did not alter their global gene expression levels or the expression
335 profiles of seven sensitive biological markers.
336

337 **FUNDING STATEMENT**

338 This work was supported by the National GMO New Variety Breeding Program of PRC
339 (2014ZX0801102B-003), the National Natural Sciences Foundation of China (31501849,
340 31471958, 31772510) and the Open Project of State Key Laboratory for Biology of Plant
341 Diseases and Insect Pests (SKLOF201306).

342 **REFERENCES**

343 Baillon, L.; Pierron, F.; Pannetier, P.; Normandeau, E.; Couture, P.; Labadie, P.; Budzinski, H.;
344 Lambert, P.; Bernatchez, L.; Baudrimont, M. Gene transcription profiling in wild and
345 laboratory-exposed eels: Effect of captivity and in situ chronic exposure to pollution. *Sci.*
346 *Total Environ.* **2016**, *57*, 92—102.
347 Bakonyi, G.; Dolezsai, A.; Mátrai, N.; Székács, A. Effects of consumption of Bt-maize (*MON*
348 *810*) on the collembolan *Folsomia candida*, over multiple generations: A laboratory study.
349 *Insects.* **2011**, *2*(2), 243—252.
350 Buch, A. C.; Niemeyer, J.C.; Correia, M. E. F.; Silva, E. V. Ecotoxicity of mercury to *Folsomia*
351 *candida* and *Proisotoma minuta* (Collembola: Isotomidae) in tropical soils: Baseline for
352 ecological risk assessment. *Ecotox. Environ. Safe.* **2016**, *127*, 22—29.
353 Campiche, S.; L'Arnbert, G.; Tarradellas, J.; Becker-van Slooten, K. Multigeneration effects of
354 insect growth regulators on the springtail *Folsomia candida*. *Ecotox. Environ. Safe.* **2007**,
355 *67*(2), 180—189.
356 Chen, G.; den Braver, M. W.; van Gestel, C. A. M.; van Straalen, N. M.; Roelofs, D.
357 Ecotoxicogenomic assessment of diclofenac toxicity in soil. *Environ. Pollut.* **2015**, *199*,
358 253—260.
359 Clark, B. W.; Phillips, T.A.; Coats, J. R. Environmental fate and effects of *Bacillus thuringiensis*
360 (Bt) proteins from transgenic crops: A review. *J. Agr. Food Chem.* **2005**, *53*(12),
361 4643—4653.
362 European Food Safety Authority [EFSA]. Guidance on the environmental risk assessment of

363 genetically modified plants. *EFSA J.* **2010**.

364 Gao, Y. J.; Zhu, H. J.; Chen, Y.; Li, Y. H.; Peng, Y. F.; Chen, X. P. Safety Assessment of

365 *Bacillus thuringiensis* insecticidal proteins Cry1C and Cry2A with a zebrafish

366 embryotoxicity test. *J. Agr. Food Chem.* **2018**, *66*(17), 4336—4344.

367 van Gestel, C, A, M,; Silva, C. D. L. E.; Lam, T.; Koekkoek, J. C.; Lamoree, M. H.; Verweij, R.

368 A. Multigeneration toxicity of imidacloprid and thiacloprid to *Folsomia candida*.

369 *Ecotoxicology.* **2017**, *26*(3), 1—9.

370 Higgins, L. S.; Babcock, J.; Neese, P.; Layton, R. J.; Moellenbeck, D. J.; Storer, N. Three-year

371 field monitoring of Cry1F, event DAS-Ø15Ø7–1, maize hybrids for nontarget arthropod

372 effects. *Environ. Entomol.* **2009**, *38*(1), 281—292.

373 Hopkin, S. P. *Biology of the springtails (Insecta: Collembola)*. Oxford University Press, NY,

374 **1997**.

375 Icoz, I.; Stotzky, G. Fate and effects of insect-resistant *Bt* crops in soil ecosystems. *Soil Biol.*

376 *Biochem.* **2008**, *40*, 559—586.

377 International Service for the Acquisition of Agri-biotech Applications [ISAAA]. Global Status of

378 Commercialized Biotech/GM Crops: 2016. ISAAA Brief No. 52. **2016**. ISAAA: Ithaca,

379 New York, USA.

380 Jia, H. R.; Geng, L. L.; Li, Y. H.; Wang, Q.; Diao, Q. Y.; Zhou, T.; Dai, P. L. The effects of Bt

381 Cry1le toxin on bacterial diversity in the midgut of *Apis mellifera ligustica* (Hymenoptera:

382 Apidae). *Sci. Rep.* **2016**, *6*, 24664.

383 Jiao, Y. Y.; Yang, Y.; Meissle, M.; Peng, Y. F.; Li, Y. H. Comparison of susceptibility of *Chilo*

suppressalis and *Bombyx mori* to five *Bacillus thuringiensis* proteins. *J. Invertebr. Pathol.* **2016**, *136*, 95—99.

Kim, Y. H.; Hwang, C. E.; Kim, T. S.; Lee, J. H.; Lee, S. H. Assessment of potential impacts due to unintentionally released Bt maize plants on non-target aphid *Rhopalosiphum padi* (Hemiptera: Aphididae). *J. Asia-Pac. Entomol.* **2012**, *15*(3), 443—446.

Li, Y. H.; Romeis, J.; Wu, K. M.; Peng, Y. F. Tier-1 assays for assessing the toxicity of insecticidal proteins produced by genetically engineered plants to non-target arthropods. *Insect Sci.* **2014**, *21*(2), 125—134.

Li, Y. H.; Wu, K. M.; Zhang, Y. J.; Yuan, G. H. Degradation of Cry1Ac protein within transgenic *Bacillus thuringiensis* rice tissues under field and laboratory conditions. *Environ. Entomol.* **2007**, *36*(5), 1275—1282.

Liu, Q. S.; Hallerman, E.; Peng, Y. F.; Li, Y. H. Development of *Bt* rice and *Bt* maize in China and their efficacy in target pest control. *Int. J. Mol. Sci.* **2016**, *17*(10), 1561.

Liu, Y. R.; Zheng, Y. M.; Zhang, L. M.; Luan, Y. X.; He, J. Z. Effects of mercury on reproduction, avoidance, and heat shock protein gene expression of the soil springtail *Folsomia candida*. *Environ. Toxicol. Chem.* **2010**, *30*(7), 654—659.

Marques, L. H.; Santos, A. C.; Castro, B. A.; Storer, N. P.; Babcock, J. M.; Lepping, M. D.; Sa, V.; Moscardini, V. F.; Rule, D. M.; Fernandes, O. A. Impact of transgenic soybean expressing Cry1Ac and Cry1F proteins on the non-target arthropod community associated with soybean in Brazil. *Plos One.* **2018**, *13*(2), e0191567.

Montgomery, S. B.; Sammeth, M.; Gutierrez-Arcelus, M.; Lach, R. P.; Ingle, C.; Nisbett, J.;

405 Guigo, R.; Dermitzakis, E. T. Transcriptome genetics using second generation sequencing
406 in a Caucasian population. *Nature.* **2010**, *464*(7289), 773—777.
407 Nakamori, T.; Kaneko, N. Biomarker responses reveal that food quality affects cadmium
408 exposure in the soil collembolan *Folsomia candida*. *Environ. Pollut.* **2013**, *176*, 165—170.
409 Niu, L.; Ma, W. H.; Lei, C. L.; Jurat-Fuentes, J. L.; Chen, L. Z. Herbicide and insect resistant Bt
410 cotton pollen assessment finds no detrimental effects on adult honey bees. *Environ. Pollut.*
411 **2017**, *230*, 479—485.
412 Organisation for Economic Cooperation and Development [OECD]. *Test No. 232: Collembolan*
413 *Reproduction Test in Soil*. OECD Publishing, Paris, France, **2016**.
414 Oliveira, M.; Cardoso, D. N.; Soares, A. M. V. M.; Loureiro, S. Effects of short-term exposure to
415 fluoxetine and carbamazepine to the collembolan *Folsomia candida*. *Chemosphere.* **2015**,
416 *120*, 86—91.
417 Qiao, M.; Wang, G. P.; Zhang, C.; Roelofs, D.; van Straalen, N. M.; Zhu, Y. G. Transcriptional
418 profiling of the soil invertebrate *Folsomia candida* in pentachlorophenol-contaminated soil.
419 *Environ. Toxicol. Chem.* **2015**, *34*(6), 1362—1368.
420 Roelofs, D.; Janssens, T. K. S.; Timmermans, M. J. T. N.; Nota, B.; Mariën, J.; Bochdanovits,
421 Z.; Ylstra, B.; van Straalen, N. M. Adaptive differences in gene expression associated with
422 heavy metal tolerance in the soil arthropod *Orchesella cincta*. *Mol. Ecol.* **2009**, *18*(15),
423 3227—3239.
424 Romeis, J.; Bartsch, D.; Bigler, F.; Candolfi, M. P.; Gielkens, M. M. C.; Hartley, S. E.;
425 Hellmich, R. L.; Huesing, J. E.; Jepson, P. C.; Layton, R.; Quemada, H.; Raybould, A.;

426 Rose, R. I.; Schiemann, J., Sears, M. K.; Shelton, A. M.; Sweet, J.; Vaituzis, Z.; Wolt, J. D.
427 Assessment of risk of insect-resistant transgenic crops to nontarget arthropods. *Nat.*
428 *Biotechnol.* **2008**, *26*(2), 203—208.
429 Shelton, A. M.; Zhao, J. Z.; Roush, R. T. Economic, ecological, food safety, and social
430 consequences of the deployment of Bt transgenic plants. *Annu. Rev. Entomol.* **2002**, *47*,
431 845—881.
432 Sims, S. R.; Martin, J. W. Effect of the *Bacillus thuringiensis* insecticidal proteins CryIA(b),
433 CryIA(c), CryIIA, and CryIIIA on *Folsomia candida* and *Xenylla grisea* (Insecta:
434 Collembola). *Pedobiologia.* **1997**, *41*(5), 412—416.
435 Sultan, M.; Schulz, M. H.; Richard, H.; Magen, A.; Klingenhoff, A.; Scherf, M.; Seifert, M.;
436 Borodina, T.; Soldatov, A.; Parkhomchuk, D.; Schmidt, D.; O'Keeffe, S.; Haas, S.;
437 Vingron, M.; Lehrach, H.; Yaspo, M. L. A global view of gene activity and alternative
438 splicing by deep sequencing of the human transcriptome. *Science.* **2008**, *321*(5891),
439 956—960.
440 Szabó, B.; Seres, A.; Bakonyi, G. Long-term consumption and food replacement of near-isogenic
441 by Bt-maize alter life-history traits of *Folsomia candida* Willem 1902 (Collembola). *Appl.*
442 *Ecol. Env. Res.* **2017**, *15*(4), 1275—1286.
443 Tian, J. C.; Wang, X. P.; Long, L. P.; Romeis, J.; Naranjo, S. E.; Hellmich, R. L.; Shelton, A. M.
444 Eliminating host-mediated effects demonstrates Bt maize producing Cry1F has no adverse
445 effects on the parasitoid *Cotesia marginiventris*. *Transgenic Res.* **2014**, *23*(2), 257—264.
446 Valldor, P.; Miethling-Graff, R.; Martens, R.; Tebbe, C. C. Fate of the insecticidal Cry1Ab

protein of GM crops in two agricultural soils as revealed by ^{14}C-tracer studies. *Appl. Microbiol. Biot.* **2015**, *99*(17), 7333—7341.

Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*(1), 57—63.

Xu, E. G.; Mager, E. M.; Grosell, M.; Hazard, E. S.; Hardiman, G.; Schlenk, D. Novel transcriptome assembly and comparative toxicity pathway analysis in mahi-mahi (*Coryphaena hippurus*) embryos and larvae exposed to Deepwater Horizon oil. *Sci. Rep.* **2017**, *7*, 44546.

Yang, Y.; Chen, X.P.; Cheng, L. S.; Cao, F. Q.; Romeis, J.; Li, Y. H.; Peng, Y. F. Toxicological and biochemical analyses demonstrate no toxic effect of Cry1C and Cry2A to *Folsomia candida*. *Sci. Rep.* **2015**, *5*, 15619.

Yang, Y.; Zhang, B.; Zhou, X.; Romeis, J.; Peng, Y. F.; Li, Y. H. Toxicological and biochemical analyses demonstrate the absence of lethal or sublethal effects of *cry1C*- or *cry2A*-expressing *Bt* rice on the collembolan *Folsomia candida*. *Front. Plant Sci.* **2018**, *9*, 131.

Yousef, H. A.; Abdelfattah, E. A.; Augustyniak, M. Evaluation of oxidative stress biomarkers in *Aiolopus thalassinus* (Orthoptera: Acrididae) collected from areas polluted by the fertilizer industry. *Ecotoxicology.* **2017**, *26*(3), 340—350.

Yuan, Y. Y.; Ke, X.; Chen, F. J.; Krogh, P. H.; Ge, F. Decrease in catalase activity of *Folsomia candida* fed a *Bt* rice diet. *Environ. Pollut.* **2011**, *159*(12), 3714—3720.

Yuan, Y. Y.; Krogh, P. H.; Bai, X.; Roelofs, D.; Chen, F. J.; Zhu-Salzman, K.; Liang, Y. Y.; Sun, Y. C.; Ge, F. Microarray detection and qPCR screening of potential biomarkers of

Folsomia candida (Collembola: Isotomidae) exposed to *Bt* proteins (Cry1Ab and Cry1Ac). *Environ. Pollut.* **2014**, *184*, 170—178.

Zhang, B.; Yang, Y.; Zhou, X.; Shen, P.; Peng, Y. F.; Li, Y. H. A laboratory assessment of the potential effect of Cry1Ab/Cry2Aj-containing Bt maize pollen on *Folsomia candida* by toxicological and biochemical analyses. *Environ. Pollut.* **2017**a, *222*, 94—100.

Zhang, X. H.; Ruan, Z. Q.; You, X. X.; Wang, J. T.; Chen, J. M.; Peng, C.; Shi, Q. *De novo* assembly and comparative transcriptome analysis of the foot from Chinese green mussel (*Perna viridis*) in response to cadmium stimulation. *Plos One.* **2017**b, *12*(5), e0176677.

Zortéa, T.; Baretta, D.; Maccari, A. P.; Segat, J. C.; Boiago, E. S.; Sousa, J. P.; Da Silva, A. S. Influence of cypermethrin on avoidance behavior, survival and reproduction of *Folsomia candida* in soil. *Chemosphere.* **2015**, *122*, 94—98.

Figure 1: Culture of *Folsomia candida* and the flow chart of the experiment. (A) *F. candida* reared in a Petri dish with baker's yeast as food. (B) A living individual of *F. candida*. (C) A cluster of *F. candida* eggs. (D) The flow chart of the dietary exposure experiment over three consecutive generations. The CT set of collembolans was fed yeast powder without Cry1F, and the BT set was fed yeast powder with Cry1F. In each generation, both CT and BT sets were fed and evaluated individually (with 30 or 40 replicates) or in groups of 60 to 80 individuals (with two or three replicates).

Figure 2: The reproduction (number of juveniles produced per individually fed collembolan) of three generations of *F. candida* as affected by addition of Cry1F to an artificial diet. The collembolans were individually fed pure yeast powder (CT set) or yeast powder + Cry1F protein (BT set). Values are means ± SE (n = 24-38). Across all generations, independent Student's *t*-test between CT and BT sets for each generation indicate no significant difference, with $p = 0.912$, 0.114, 0.071 for the 1st, 2nd and 3rd generations, respectively (significant difference $p < 0.05$).

Figure 3: DEGs in *F. candida* fed diets with and without Cry1F over three generations. (A-C) Volcano plots of DEGs for the three generations. DEGs: FC (|log2 ratio| of FPKM (BT vs. CT)) ≥ 1, FDR ≤ 0.05. (D) Numbers of up-regulated and down-regulated DEGs. (E) Venn diagrams of DEGs in three generations. DEGs = differentially expressed genes; FC = fold change; FPKM = fragments per kb per million fragments; FDR = false discovery rate.

500 **Figure 4:** Heatmap and hierarchical clustering dendrogram of DEGs. (A) Heatmap of DEGs.
501 Gene clusters were subjected to Gene Ontology analysis. Columns represent samples. The scale
502 bar indicates log-transformed (common logarithm of the counts plus 1) gene expression values,
503 with high expression depicted in red and low expression in green. (B) Hierarchical clustering
504 dendrogram of six samples over three generation. Log-transformed data of DEGs were used for
505 analysis. The distance method was set as 'euclidean' and the cluster method was set as
506 'complete'.
507
508 **Figure 5:** Heatmap of seven biomarkers. Gene clusters from hierarchical classification were
509 subjected to Gene Ontology analysis. Columns represent samples. The scale bar indicates log-
510 transformed gene expression values, with high expression depicted in red and low expression in
511 green.

Table 1(on next page)

A two-way ANOVA of reproduction[a] with diet treatment, generation, and their interaction as fixed factors, followed by LSD test.

Table 1. A two-way ANOVA of reproduction[a] with diet treatment, generation, and their interaction as fixed factors, followed by LSD test.

source	SS	*df*	MS	*F*	*p*-value
diet treatment	0.052	1	0.052	3.363	0.068
generation	0.227	2	0.114	7.414	0.001*
diet treatment × generation	0.027	2	0.014	0.896	0.410
error	2.956	193	0.015		
total	348.507	199			

[a] The number of juveniles produced per collembolan in the individual feeding test was calculated as reproduction value.

* $p < 0.05$.

Table 2(on next page)

Summary of *F. candida* transcriptome sequencing and assembly for CT and BT sets over three generations.

Table 2. Summary of *F. candida* transcriptome sequencing and assembly for CT and BT sets over three generations.

statistic	CT			BT		
	1st	2nd	3rd	1st	2nd	3rd
raw reads	55,491,201	52,314,588	51,127,722	55,080,974	49,522,385	53,156,777
clean reads	51,147,377	43,971,149	47,089,957	50,778,618	42,728,374	48,458,947
Q20 (%)	98.21	98.46	99.00	98.33	98.36	98.80
GC content (%)	40.63	42.93	42.96	41.97	43.63	43.84

Table 3(on next page)

Summary of annotated unigenes and DEGs[a] of seven biomarkers of *F. candida* over three generations.

Table 3. Summary of annotated unigenes and DEGs[a] of seven biomarkers of *F. candida* over three generations.

biomarker	number of unigenes	name of unigene	1st		2nd		3rd	
			FC[b]	FDR	FC	FDR	FC	FDR
CAT	52	CL2174.Contig5_All	0.443	1.000	1.859	1.000	2.857	0.042*
GST	191	-	-	-	-	-	-	-
SOD	42	-	-	-	-	-	-	-
GR	6	-	-	-	-	-	-	-
		Unigene7933_All	1.664	0.042*	-2.469	1.000	0.225	1.000
CES	408	CL5034.Contig2_All	-1.968	0.023*	0.366	1.000	-0.410	1.000
		CL6365.Contig1_All	0.794	1.000	-0.104	1.000	2.054	0.000*
MTC	38	-	-	-	-	-	-	-
HSP70	118	Unigene15285_All	-5.226	0.043*	-2.191	1.000	-2.341	1.000
SUM	855							

[a] DEGs: $|FC| \geq 1$, FDR ≤ 0.05.

[b] FC: log2 ratio of FPKM (BT vs. CT).

* FDR ≤ 0.05.

DEGs = differentially expressed genes; FC = fold change; FDR = false discovery rate; FPKM = fragments per kb per million fragments; CAT = catalase; GST = glutathione S-transferase; SOD = superoxide dismutase; GR = glutathione reductase; CES = carboxylesterase; MTC = metallothionein-like motif containing protein; HSP70 = heat shock protein 70.

Figure 1

Culture of *Folsomia candida* and the flow chart of the experiment.

(A) *F. candida* reared in a Petri dish with baker's yeast as food. (B) A living individual of *F. candida*. (C) A cluster of *F. candida* eggs. (D) The flow chart of the dietary exposure experiment over three consecutive generations. The CT set of collembolans was fed yeast powder without Cry1F, and the BT set was fed yeast powder with Cry1F. In each generation, both CT and BT sets were fed and evaluated individually (with 30 or 40 replicates) or in groups of 60 to 80 individuals (with two or three replicates).

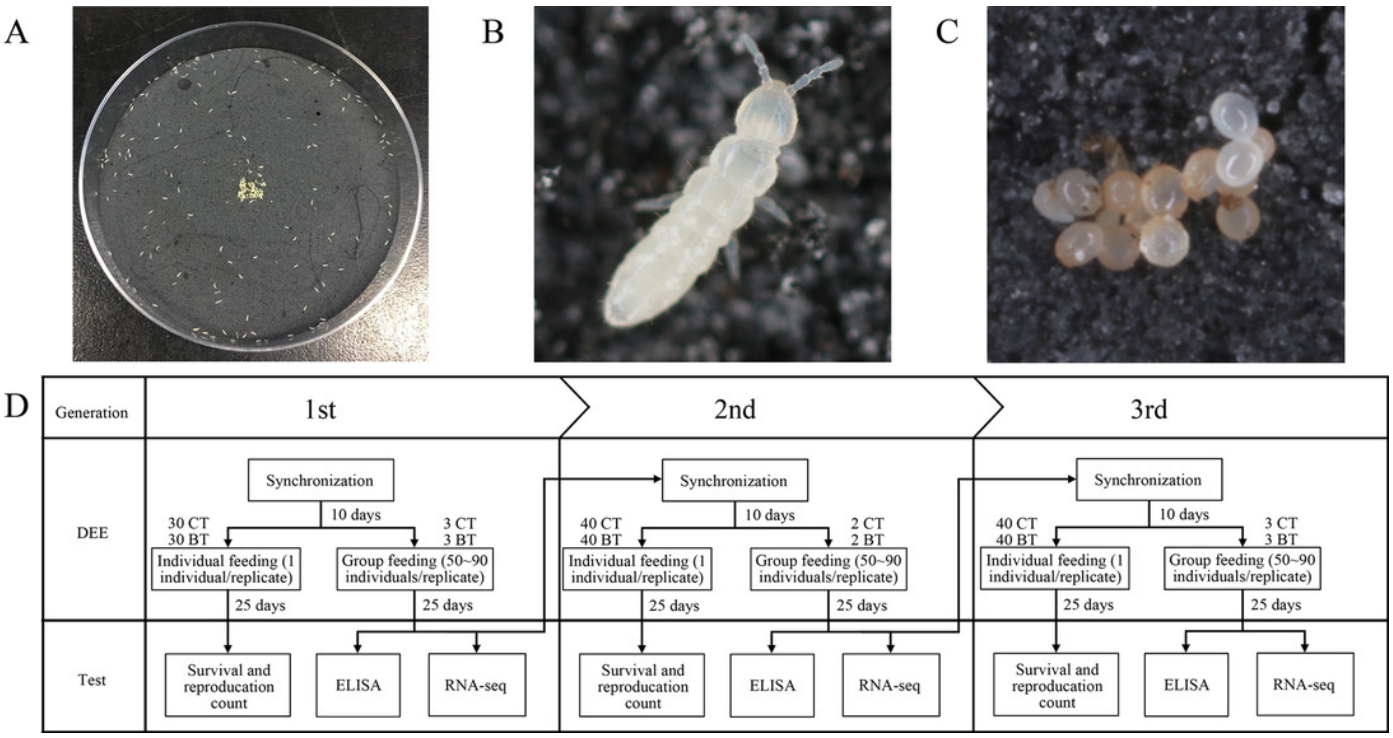

Figure 2

The reproduction (number of juveniles produced per individually fed collembolan) of three generations of *F. candida* as affected by addition of Cry1F to an artificial diet.

The collembolans were individually fed pure yeast powder (CT set) or yeast powder + Cry1F protein (BT set). Values are means ± SE (n = 24-38). Across all generations, independent Student's *t*-test between CT and BT sets for each generation indicate no significant difference, with $p = 0.912, 0.114, 0.071$ for the 1st, 2nd and 3rd generations, respectively (significant difference $p < 0.05$).

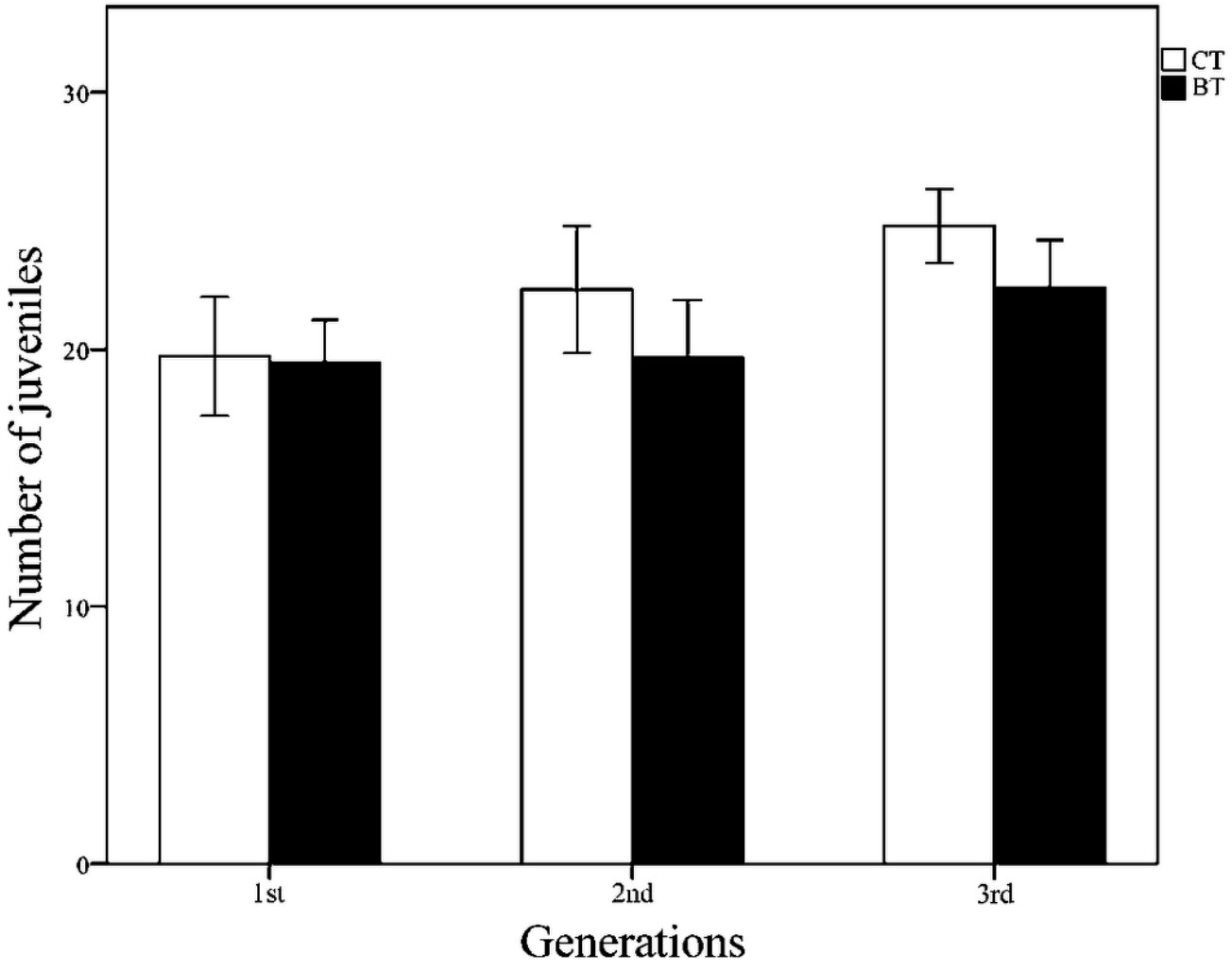

Figure 3

DEGs in *F. candida* fed diets with and without Cry1F over three generations.

(A-C) Volcano plots of DEGs for the three generations. DEGs: FC (|log2 ratio| of FPKM (BT vs. CT)) ≥ 1, FDR ≤ 0.05. (D) Numbers of up-regulated and down-regulated DEGs. (E) Venn diagrams of DEGs in three generations. DEGs = differentially expressed genes; FC = fold change; FPKM = fragments per kb per million fragments; FDR = false discovery rate.

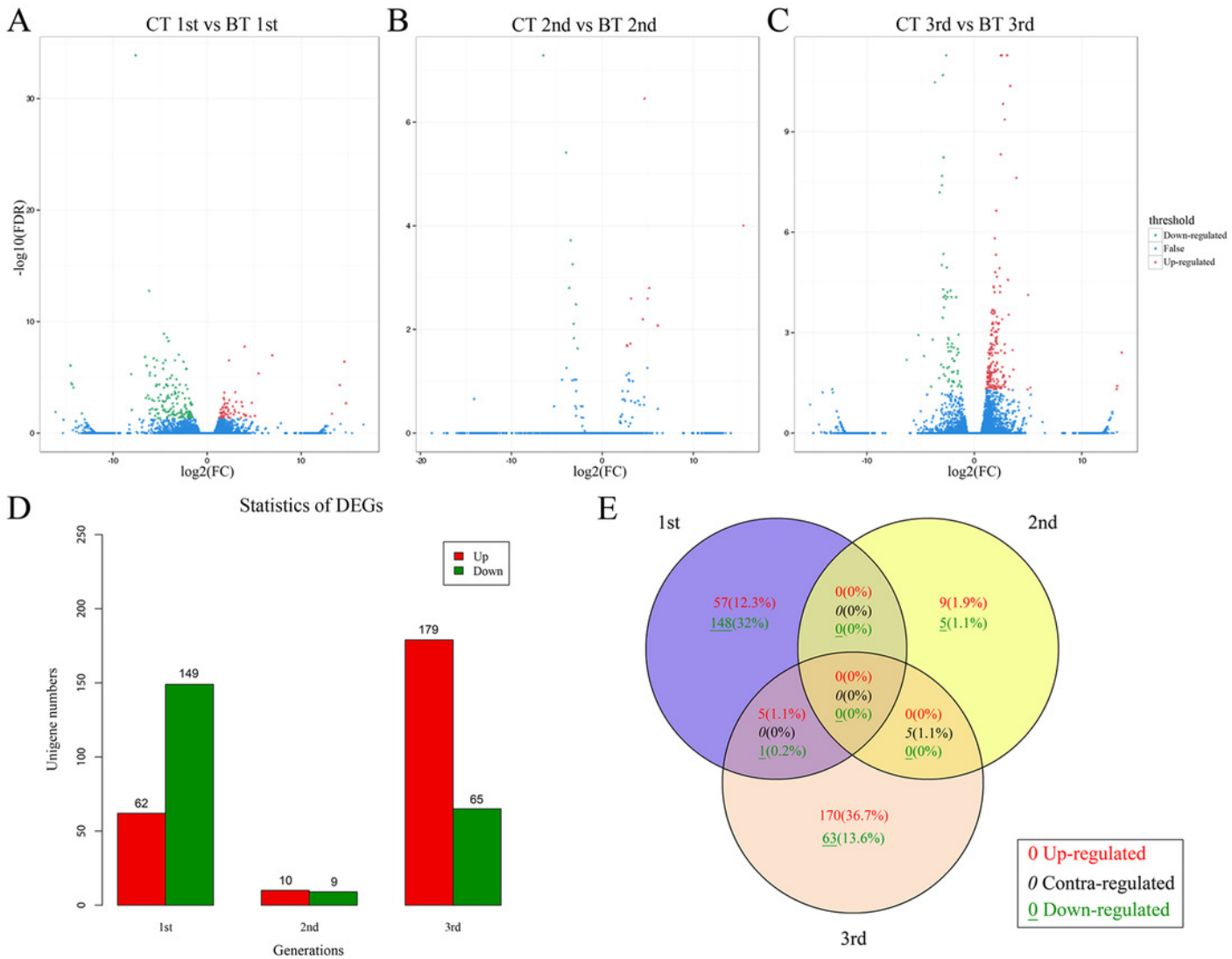

Figure 4

Heatmap and hierarchical clustering dendrogram of DEGs.

(A) Heatmap of DEGs. Gene clusters were subjected to Gene Ontology analysis. Columns represent samples. The scale bar indicates log-transformed (common logarithm of the counts plus 1) gene expression values, with high expression depicted in red and low expression in green. (B) Hierarchical clustering d endrogram of six samples over three generation. Log-transformed data of DEGs were used for analysis. The distance method was set as ‘euclidean’ and the cluster method was set as ‘complete’.

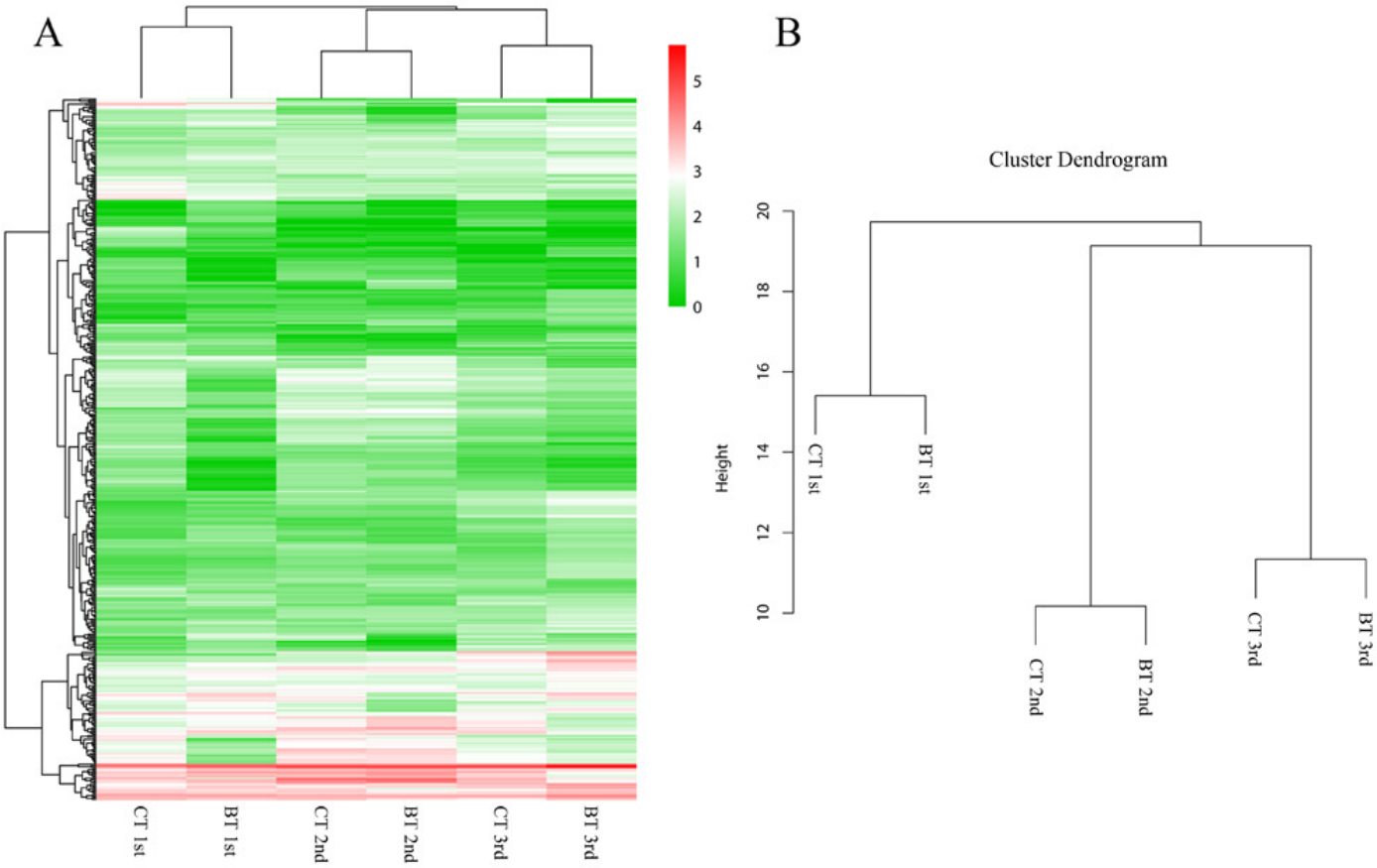

Figure 5

Heatmap of seven biomarkers.

Gene clusters from hierarchical classification were subjected to Gene Ontology analysis. Columns represent samples. The scale bar indicates log-transformed gene expression values, with high expression depicted in red and low expression in green.