Genome-wide bisulfite sequencing data of normal and *Ascosphaera apis*-infected larval guts of eastern honeybee

Yu Du[#], Zhiwei Zhu[#], Huazhi Chen, Yuanchan Fan, Jie Wang, Xiaoxue Fan, Haibin Jiang, Cuiling Xiong, Yanzhen Zheng, Dafu Chen^{*}, Rui Guo^{*}

College of Animal Sciences (College of Bee Science), Fujian Agriculture and Forestry University, Fuzhou 350002, China

^{*} Corresponding author: Tel: +86-0591-87640197; Fax: +86-0591-87640197; E-mail address: dfchen826@fafu.edu.cn; ruiguo@fafu.edu.cn

ABSTRACT

Apis cerana cerana, a subspecies of eastern honey, *Apis cerana*, plays a specific role in beekeeping industry and ecosystem in China and other Asian countries. Larvae of *A. c. cerana* can be infected by *Ascosphaera apis*, the fungal pathogen of chalkbrood. In this article, normal 4-, 5-, and 6-day-old larval guts (AcCK1, AcCK2, AcCK3) and *A. apis*-infected 4-, 5- and 6-day-old larval guts (AcT1, AcT2, AcT3) of *A. c. cerana* workers were respectively harvested followed by DNA isolation, bisulfite conversion, cDNA library construction and Illumina sequencing. Based on genome-wide bisulfite sequencing, 79167210, 82175052, 79331489, 81051009, 74742842 and 74849091 raw reads were generated from AcCK1, AcCK2, AcCK3, AcT1, AcT2 and AcT3, and after quality control, 73417030 (92.73%), 76660370 (93.27%), 71804727 (90.44%), 75046507 (92.82%), 67487782 (90.30%) and 67367023 (90.04%) clean reads were obtained, respectively. Additionally, 73333333, 76533333, 71466667, 75066667, 67590965 and 67200000 clean reads were mapped to the reference genome of *A. cerana*, including 54656767, 58583415, 54127407, 57943220, 52547867 and 51295824 unique mapped clean reads, and 8624392, 8789458, 7531333, 7747337, 6249679 and 5394174 multiple mapped clean reads. The genome-wide bisulfite sequencing data reported here can be used for genome-wide

identification of 5mC methylation sites in eastern honeybee larval guts and systematic investigation of DNA methylation-mediated host response to *A. apis* infection.

Keywords

Genome-wide bisulfite sequencing, Methylome, Methylation, Honeybee, Chalkbrood, *Ascosphaera apis*

Value of the data

- The current dataset contributes to genome-wide identification of 5mC methylation sites in normal and *A. apis*-infected larval guts of eastern honeybee.
 - The reported data could be used for systematic investigation of DNA methylation-mediated response of eastern honeybee larvae to *A. apis* infection.
 - Our data offers a valuable genetic resource for better understanding epigenetic regulation mechanism involved in eastern honeybee larvae-*A. apis* interaction.
-

1. Data description

The data reported here were obtained from genome-wide bisulfite sequencing of normal and *Ascosphaera apis*-infected larval guts of *Apis cerana cerana*. The ratio of G and C among total bases derived from genome-wide bisulfite sequencing of normal 4-, 5-, and 6-day-old larval guts (AcCK1, AcCK2, AcCK3) and *A. apis*-infected 4-, 5-, and 6-day-old larval guts (AcT1, AcT2, AcT3) was obviously much lower than that of A and T (**Figure 1, Figure 2, Figure 3, Figure 4, Figure 5, Figure 6**). In total, 79167210, 82175052, 79331489, 81051009, 74742842 and 74849091 raw reads were yielded from AcCK1, AcCK2, AcCK3, AcT1, AcT2 and AcT3 groups, amounting to about 11.88 Gb, 12.33 Gb, 11.90 Gb, 12.16 Gb, 11.21 Gb and 11.23 Gb, respectively (**Table 1**). After quality control, 73417030 (92.73%), 76660370 (93.27%), 71804727

(90.44%), 75046507 (92.82%), 67487782 (90.30%) and 67367023 (90.04%) clean reads were gained, amounting to about 10.99 Gb, 11.48 Gb, 10.75 Gb, 11.24 Gb, 10.11 Gb and 10.09 Gb, respectively (**Table 1**). Additionally, the GC contents of clean reads from the aforementioned six groups were 17.33%, 17.33%, 17.83%, 17.67%, 18.17% and 19%, respectively (**Table 1**). Furthermore, 73333333, 76533333, 71466667, 75066667, 67590965 and 67200000 clean reads were mapped to the reference genome of *Apis cerana*, including 54656767, 58583415, 54127407, 57943220, 52547867 and 51295824 unique mapped clean reads, and 8624392, 8789458, 7531333, 7747337, 6249679 and 5394174 multiple mapped clean reads (**Table 2**).

Figure 1 Ratio of four different bases from normal 4-day-old larval gut of *A. c. cerana*. The x axis indicates the position in reads, and the y axis indicates the ratio of a single base among total bases.

Figure 2 Ratio of four different bases from normal 5-day-old larval gut of *A. c. cerana*. The x axis indicates the position in reads, and the y axis indicates the ratio of a single base among total bases.

Figure 3 Ratio of four different bases from normal 6-day-old larval gut of *A. c. cerana*. The *x* axis indicates the position in reads, and the *y* axis indicates the ratio of a single base among total bases.

Figure 4 Ratio of four different bases from *A. apis*-infected 4-day-old larval gut of *A. c. cerana*. The *x* axis indicates the position in reads, and the *y* axis indicates the ratio of a single base among total bases.

Figure 5 Ratio of four different bases from *A. apis*-infected 5-day-old larval gut of *A. c. cerana*. The *x* axis indicates the position in reads, and the *y* axis indicates the ratio of a single base among total bases.

Figure 6 Ratio of four different bases from *A. apis*-infected 6-day-old larval gut of *A. c. cerana*. The *x* axis indicates the position in reads, and the *y* axis indicates the ratio of a single base among total bases.

Table 1 Summary of genome-wide bisulfite sequencing of *A. c. cerana* larval guts in control and treatment groups.

Samples	Raw reads	Clean reads	Percentage of clean reads	Raw bases (bp)	Clean bases (bp)	Percentage of clean bases	GC content (%)
AcCK1-1	69418114	64057732	92.28%	10412717100	9593136522	92.13%	18.00
AcCK1-2	95041508	88178256	92.78%	14256226200	13205415336	92.63%	17.00
AcCK1-3	73042008	68015102	93.12%	10956301200	10185737274	92.97%	17.00
AcCK2-1	79995726	74864346	93.59%	11999358900	11211755592	93.44%	17.00
AcCK2-2	76949428	71301404	92.66%	11542414200	10677515702	92.51%	17.00
AcCK2-3	89580002	83815360	93.56%	13437000300	12553243137	93.42%	18.00
AcCK3-1	78565504	73656682	93.75%	11784825600	11031899396	93.61%	16.50
AcCK3-2	76054260	65507216	86.13%	11408139000	9811214511	86.00%	19.00
AcCK3-3	83374702	76250282	91.45%	12506205300	11420719136	91.32%	18.00
AcT1-1	69369854	65498754	94.42%	10405478100	9810143035	94.28%	18.00
AcT1-2	95841572	86676046	90.44%	14376235800	12982823947	90.31%	17.00
AcT1-3	77941602	72964722	93.61%	11691240300	10927742957	93.47%	18.00
AcT2-1	72210136	64665250	89.55%	10831520400	9683970609	89.41%	18.00
AcT2-2	71943768	65707650	91.33%	10791565200	9841409109	91.20%	18.00
AcT2-3	80074622	72090446	90.03%	12011193300	10797230783	89.89%	18.50
AcT3-1	69055652	62557190	90.59%	10358347800	9369980507	90.46%	18.00
AcT3-2	87329208	78254536	89.61%	13099381200	11711973583	89.41%	20.00
AcT3-3	68162412	61289344	89.92%	10224361800	9178648933	89.77%	19.00

Table 2 Summary of mapping of clean reads to the reference genome of *A. cerana*.

Samples	Clean reads	Unique mapped reads	Multiple mapped reads	Unmapped reads	Mapping ratio
AcCK1-1	64000000	48299599	7244955	8455446	86.79%
AcCK1-2	88000000	63877814	11025730	13096456	85.12%
AcCK1-3	68000000	51792887	7602491	8604622	87.35%
AcCK2-1	74800000	56188178	9396741	9215081	87.68%
AcCK2-2	71200000	53830285	8624498	8745217	87.72%
AcCK2-3	83600000	65731781	8347135	9521084	88.61%
AcCK3-1	73200000	55699500	9611064	7889436	89.22%
AcCK3-2	65200000	50506793	5305352	9387855	85.60%
AcCK3-3	76000000	56175929	7677583	12146488	84.02%
AcT1-1	66000000	50787912	6594927	8617161	86.94%
AcT1-2	86400000	65727414	9537471	11135115	87.11%
AcT1-3	72800000	57314334	7109613	8376053	88.49%
AcT2-1	64665248	49607260	6509890	8548098	86.78%
AcT2-2	66107648	50574533	6014641	9518474	85.60%
AcT2-3	72000000	57461807	6224507	8313686	88.45%
AcT3-1	62400000	49484583	5664547	7250870	88.38%
AcT3-2	78000000	56976725	5996043	15027232	80.73%
AcT3-3	61200000	47426164	4521931	9251905	84.88%

2. Experimental Design, Materials, and Methods

2.1. Preparation of eastern honeybee larval gut samples

Two-day-old larvae were removed from the combs with a Chinese grafting tool and seriously transferred to a droplet of 10 μ L diet. Larvae were feed once a day with 20 μ L (3-day-old), 30 μ L (4-day-old), 40 μ L (5-day-old) and 50 μ L (6-day-old) diet, 150 μ L diet total per larva. Honeybee larvae were reared in 48-cell culture plates according to the method described by Peng et al. [1].

A. apis [2] was previously isolated and conserved at the Honeybee Protection Laboratory of the College of Animal Sciences (College of Bee Science) in Fujian Agriculture and Forestry University. Before experimental inoculation, *A. apis* was

cultured at 33 ± 0.5 °C on plates of Potato dextrose agar (PDA) medium followed by purification of fungal spores using the method developed by Jensen et al. [3].

Three-day-old larvae in treatment groups were reared with artificial diet containing *A. apis* spores (1×10^7 spores/mL), while 3-day-old larvae in control groups were reared with artificial diet without fungal spores. Then, 4-, 5- and 6-day-old larvae in both control groups and treatment groups were reared with artificial diet without *A. apis* spores. Using the method previously developed in our lab [4], guts of 4-, 5- or 6-day-old larvae from *A. apis*-treated groups (n=21) and normal groups (n=21) were respectively harvested.

2.2. Bisulfite conversion and library preparation

Firstly, genomic DNA was respectively extracted from AcCK1, AcCK2, AcCK3, AcT1, AcT2, and AcT3 groups using a Universal Genomic DNA Extraction Kit (TaKaRa, Tokyo, Japan) according to the manufacturer's instruction. Next, DNA was bisulfite treated using a Zymo Research EZ DNA methylaiton-Glod Kits (Zymo Research, Irvine, CA, USA). DNA concentration and integrity were examined by a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and agarose gel electrophoresis, respectively. Finally, the eighteen libraries were constructed with TruSeq DNA Methylation Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol.

2.3. Next-generation sequencing and data processing

The constructed eighteen cDNA libraries were sequenced on an Illumina HiSeq X Ten platform by OE Biotech Co., Ltd. (Shanghai, China) and 150 bp paired-end reads were generated. After being processed with fastp [5], clean reads were gained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw reads, and used for downstream analyses. Further, the clean reads were aligned to the *A. cerana* genome (assembly ACSNU-2.0) using the genome-wide bisulfite sequencing mapping program (BSMAP, v2.74) [6]. Mapping ratio was then calculated for each sample group.

Acknowledgments

This research was financially supported by the Earmarked Fund for China Agriculture Research System (No. CARS-44-KXJ7), the Science and Technology Planning Project of Fujian Province (No. 2018J05042), the Teaching and Scientific Research Fund of Education Department of Fujian Province (No. JAT170158), the Outstanding Scientific Research Manpower Fund of Fujian Agriculture and Forestry University (No. xjq201814), and the Scientific and Technical Innovation Fund of Fujian Agriculture and Forestry University (No. CXZX2017342, No. CXZX2017343).

Conflict of interest

The authors declare that they have no competing financial interests.

Reference

- [1] Y.S.C. Peng, E. Mussen, A. Fong, M.A. Montague, T.Tyler. Effects of chlortetracycline of honey bee worker larvae reared *in vitro*. J. Invertebr. Pathol. (1992), 60(2): 127-133. doi: [http://doi.org/10.1016/0022-2011\(92\)90085-I](http://doi.org/10.1016/0022-2011(92)90085-I).
- [2] R. Guo, D.F. Chen, H.Z. Chen, Z.M. Fu, C.L. Xiong, C.S. Hou, Y.Z. Zheng, Y.L. Guo, H.P. Wang, Y. Du, Q.Y. Diao. Systematic investigation of circular RNAs in *Ascosphaera apis*, a fungal pathogen of honeybee larvae. Gene. (2018), 678: 17-22. doi: <http://doi.org/10.1016/j.gene.2018.07.076>.
- [3] A.B. Jensen, K. Aronstein, J.M. Flores, S. Vojvodic, M.A. Palacio, M. Spivak. Standard methods for fungal brood disease research. J. Apic. Res. (2013), 52(1), doi: <http://doi.org/10.3896/IBRA.1.52.1.13>.
- [4] R. Guo, D.F. Chen, Q.Y. Diao, C.L. Xiong, Y.Z. Zheng, C.S. Hou. Transcriptomic investigation of immune responses of the *Apis cerana cerana* larval gut infected by *Ascosphaera apis*. J. Invertebr. Pathol. (2019), 166: 107210. doi: <http://doi.org/10.1016/j.jip.2019.107210>
- [5] S. Chen, Y. Zhou, Y. Chen, J. Gu. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 2018, 34(17): i884-i890, doi: <http://doi.org/10.1093/bioinformatics/bty560>.
- [6] Y. Xi, W. Li. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinf. 2009, 10, 232, doi: <http://doi.org/10.1186/1471-2105-10-232>.