Genome Size Survey of the Japanese Pine Sawyer Beetle, *Monochamus alternatus* (Coleoptera: Cerambycidae)

Bo Fan [a], Fang Hong [a], Wen-Jin Su [b], Guo-Fang Jiang [a*] and Jing-Hui Zhong [b]

[a] *College of Oceanology and Food Science, Quanzhou Normal University, Quanzhou-362000, Fujian, China.*
[b] *Quanzhou Station of Forestry Pest Control and Quarantine, Quanzhou Forestry Bureau, Quanzhou-362000, Fujian, China.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

Original Research Article

ABSTRACT

Monochamus alternatus is a longicorn beetle, an important stem borer of pine trees in China. In this study, we determined the first accurate genomic survey of *M. alternatus* by the next-generation sequencing (NGS), which is a relatively new methodology that can ensure the identification of large numbers of simple sequences repeat (SSR) markers, in order to increases the abundance of the Cerambycidae genome information and assist in phylogenetic, molecular systematics and evolutionary studies for Coleoptera. The result showed that the genome size was 871.09 Mb, the GC content of the genome was within normal limits (34.45%), the proportion of repetitive sequences was high (59.40%), and the heterozygosity rate was low (1.04%). The heterozygosity of *M. alternatus* is higher than 0.5% and the repeat rate is more than 50%. Based on this we inferred the use of Illumina+PacBio in sequencing assembly strategy is highly recommended.

_____

*Corresponding author: E-mail: cnjgf1208@163.com, cnjgf12082022@outlook.com;

1. INTRODUCTION

Monochamus alternatus, belonging to the subfamily Lamiinae of Coleoptera, is not only an important stem borer of pine trees in China, but also the main vector of pine wood nematode disease [1]. *M. alternatus* plays a key role in carrying, spreading and assisting the pathogen to invade the host in the process of spread and infection of pine wood nematode [2]. Natural enemies include pathogenic microorganisms, parasitic nematodes, parasitic insects, predatory insects, spiders, birds and so on. The life cycle can be divided into four stages: pupa, larva, egg and adult, with one generation a year, and the mature larva overwinters in the tunnel. Distributed in China, South Korea, Japan, Laos. It was listed on the list of forestry dangerous pests of the State Forestry Administration of China in 2003 [3]. In terms of its genetic characteristics, the whole genome of *M. alternatus* has not yet been published [4,5]. In view of the great harm of this species to forestry, we use the next-generation sequencing (NGS) technology to determine and analyze its whole genome in order to provide basic molecular data for further study of its behavior, physiology, ecology, genetic variation, evolutionary adaptation , population protection and the formulation of conservation strategies [6-11].

Despite its great harmless to pine wood, the genetic information of *M. alternatus* has remained largely unknown, the genome size and genome-wide sequencing of *M. alternatus* have not been reported. The mitogenomes of *M. alternatus, Bursaphelenchus mucronatus*, and some of the closely related species have been studied [12]. Formerly, scholars have used morphology, anatomy and physiology, and other methods to establish relationships among species for creating classification systems. However, the evolutionary relationships between *M. alternatus* and other closely related species are hard to describe just based on traditional dates. Previous studies have used mitogenome dates to assess the gene expression of *M. alternatus* [13]. Thus, we aim to expand the database of Lamiinae and provide some useful information for further study.

With the development of the sequencing techniques, more and more manuscript genomes have been assembled. The main difference between genomics and traditional biological methods lies in the scale of research, because the goal of genomics is to analyze a large

number of genes extensively and may even involve a complete set of genes that make up the genome, rather than being limited to one or a small number of genes. With the development of DNA sequence technology, the cost has been decreased rapidly and the ability of obtaining the date has been increased vastly. NGS technology has significantly increased the sequence output while reducing time and cost. Genomic survey is a method that combine NGS technology with *K*-mer dates to achieve species genome size, GC content, heterozygosity rate and repetition rate. This technology has been used widely and could also accurately predict the whole genome sizes.

In this study, we provided valuable molecular data to reveal the relationship between *M. alternatus* and pine wood nematode [14] proved that *M. alternatus* infected with *B. xylophilus* showed an increase in the expression of some antioxidant genes, so that it could obtain immune tolerance, this study showed that the existence of *B. xylophilus* would increase the expression level of some metabolic, so some scholars inferred that the relationship between nematodes and beetles should be host switching rather than nematode-beetle coevolution, because in their study, they found some efforts were made by *M. alternatus* to offset the impact. The mechanisms need to be studied in the future.

2. METHODS AND MATERIALS

During the inspection in An'xi county of Quanzhou, we had obtained the *M. alternatus* specimens, which were frozen in liquid nitrogen and stored in a cryogenic refrigerator at 80 degrees below zero.

The genomic DNA, was extracted by improved SDS method. The Absorbance values of 260 nm and 280 nm were determined by ultraviolet spectrophotometer to estimate the concentration and purity of the extracted DNA, and then agarose gel electrophoresis was performed to detect the integrity of the extracted DNA. When sampling, 1 kb DNA Ladder (Takara) and λ-Hind di-gest (Takara) were used as Marker, gel concentration of 0.8%, electrophoresis for 60 minutes under the condition of 80V voltage. After the electrophoresis was completed, the electrophoretic band of genomic DNA was observed in the gel imaging system of BIO-RAD company.

The DNA samples were randomly interrupted by Covaris ultrasonic crusher, and the small

fragment sequencing library of 250 bp was established. Then the whole library was prepared by terminal repair, a tail, sequencing joint, purification, PCR amplification and so on. After that, double-terminal (Pair-End) sequencing was carried out on the Illumina HiseqTM2000 platform. Finally, the quality control is carried out by using the software NGS-QC-Generator, the low-quality data is filtered out, and the effective data obtained are used for genome feature evaluation and preliminary assembly.

K-mer analysis was used to estimate the genome size. Use the formula: genome size = *K*-mer number / peak depth to estimate genome size. The estimated value of *K*-mer depth is calculated, which is used to estimate the genome size. The heterozygosity (Φ) was calculated by the following formula, while the repetition rate was calculated by calculating the ratio of the number of *K*-mer to the total number after 1.8 times the depth of the homozygous peak. In the formula, K is the number of heterozygous *K*-mer, $a_{1/2}$ is the percentage of heterozygous *K*-mer species, and $n_{Kspecies}$ is the number of all *K*-mer species.

The GC depth content of the target genome was calculated by constructing the GC dot map from the total number of bases in the sequencing data, and the correlation analysis was carried out. Contigs were rearranged by all clean reads and scaffolds were gradually constructed by diversified insert size paired-ends.

The sequence reads obtained from all the small fragment libraries were truncated into smaller sequence fragments, and the de-Brujin map was constructed by using the overlapping relationship between them. Then the simplified de-Brujin map which is easy to analyze was obtained by screening branches, simplifying the bifurcation pathway and randomly combining heterozygous sites, and then the bifurcation sequence was truncated into small fragments to get the initial overlapping (contigs). The reads obtained by sequencing was compared, and the obtained contigs, assembled the contigs into scaffolds by using the connection relationship between reads and the size information of insertion fragments. In order to make the assembled sequence more complete, it is necessary to connect the contigs, according to the pairing relationship between the double-terminal data and optimize and fill holes in the gaps between the contig, so as to obtain the original genome sequence.

3. RESULTS

3.1 Sequencing Quality

The low-quality data is filtered out by screening, and the effective data with the size of 34.68Gb is obtained, and the sequencing depth is 51x. Through the strict filtering of the obtained data, the high-quality pure data is obtained. The 350 bp library is constructed, the original data is 34,683,801,900bp, and the clean data 34,652,586,478bp is obtained after filtering. Both Q20 and Q30 are indicators to measure the quality of sequencing. It is generally believed that when Q20 ≥ 90% and Q30 ≥ 80%, the quality of sequencing data is better. In this study, the content of Q20 is 97.86% and the content of Q30 is 93.48%. It is known that the sequencing quality of this study is good, and the sequencing error rate is 0.04%, which is also within the normal range (< 0.05%).

After genome regulation, if the estimated heterozygosity of the genome is high (higher than 0.5%), or the content of repetitive sequences is high (more than 50%), the genome can be regarded as a complex genome. WGS sequencing and assembly strategies suitable for general genomes are generally difficult to obtain good complex genome maps. At present, there are two commonly used strategies to solve complex genome maps, SangeH-454 sequencing and WGS+BAC to BAC/fosmid to fosmid. Compared with the former, the latter is based on Illumina sequencing, and the cost is much lower. In this study, we found the heterozygosity of *M. alternatus* is higher than 0.5%, the repeat rate is more than 50% either. In consequence, the use of Illumina+PacBio in sequencing assembly strategy is highly recommended.

3.2 17-Mer Analysis and Genome Size Estimation

Genome size of *M. alternatus* was estimated according to the Lander waterman algorithm based on *K*-mer ($k = 17$) frequency of the clean reads and the 17-mer frequency distribution complied with the poisson distribution. The 34.65 Gb valid data of *M. alternatus* were analyzed by *K*-mer (Fig. 1). It can be seen that the peak is near depth = 29; the calculated genome size is 871.09 Mbp, the modified genome size is 860.06 Mbp. There is a tail on the *K*-mer curve, indicating that the proportion of genome repeat sequences of *M. alternatus* is large.
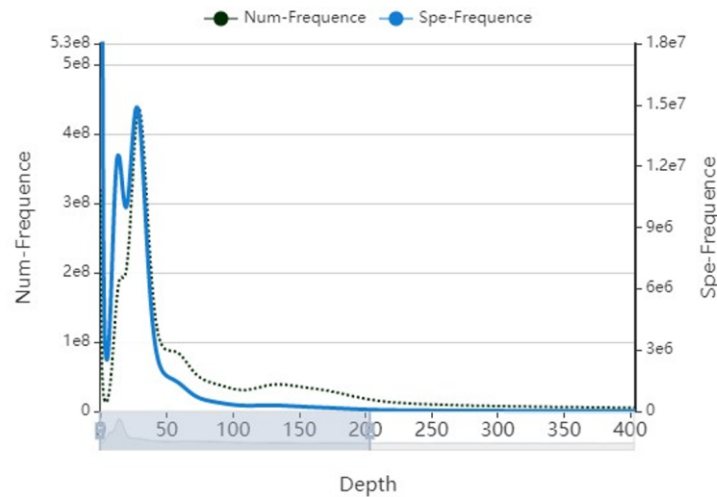
Fig. 1. 17 *K*-mer analysis for estimating the genome size of *Monochamus alternatus*
Note: The genome size was estimated by using the formula: Genome size=K-me num/Peak depth

Table 1. Statistics of assembly results in the *Monochamus alternatus* genome

Item	Contig	Scaffold
	Size (bp)	Size (bp)
N50	787	954
N90	135	145
Total Size	823 399 058	840 047 447
Total Number	2 034 890	1 839 509
Longest	385 511	458 244

3.3 Results of Preliminary Assembly of the Genome

Using SOAP-Denovo software to assemble the data, the initial genome sequence is obtained, and the splicing result is shown in Table 1. It can be seen that the genomic Contig N50 of *M. alternatus* is 787 bp and scaffolding N50 is 954 bp, which is short, so it is not suitable for further assembly by shotgun method.

3.4 Content and Distribution of GC

GC content is an important character of nucleic acid sequence composition in living creatures, which is also used to detect the separation of AT and GC. Theoretically, the contents of G and C bases and the A and T bases should be equal in each sequencing cycle, and the whole sequencing process is stable and horizontal. In the actual sequencing process, due to the DNA template amplification deviation and the low sequencing quality value of the first few bases, it often leads to large fluctuations in the first few bases of each read, which is a normal situation. According to the statistics of GC content of assembled contig, the distribution map of GC content and sequencing depth (depth) was

obtained (Fig. 2). The genomic GC content of *M. alternatus* was 32.94%.

3.5 Genome Comparison of the Insects in Coleoptera

In the genome database of the National Biotechnology Information Center of the United States (https://www.ncbi.nlm.nih.gov/genome), enter the key words "Coleoptera" to obtain 61 pieces of genome information of Coleoptera insects. We select 25 species that can be reported in the relevant literatures, and compare their genome information with the genome data of *M. alternatus* (Table 2). According to the published data, the genome sizes of Cerambycidae range from 90.87 to 1112.44 Mb, while the genome size of *M. alternatus* is as high as 871.09 Mb. Some studies believed that if the content of genomic GC is moderate (25% < GC < 65%), it will not affect the accuracy of genome sequencing and correct assembly. The content of genomic GC measured in this study is 32.94%, so the sequencing results and assembly should be correct. On the other hand, the GC contents of the other 24 genomes were similar.

4. DISCUSSION

Although the depth of our *M. alternatus* preliminary sequencing genome is relatively low and it's a disadvantage for an acquaintance of genome information, we still excavated some basic genome information and developed genome resources by using several bioinformation methods like *K*-mer and scaffold assembly methods. According to each analysis index, it is inferred that the genome of *M. alternatus* is a a highly complex genome, and the corresponding strategy can be used to assemble the genome.

There are many species of insects, about 1.8 million species are known, accounting for more than 3 beat 4 of the described animal species in the world [15-18]. In recent years, with the continuous improvement of DNA sequencing technology in the direction of automation [19], generalization and low cost, the continuous maturity of genomics research technology and the continuous updating of bioinformatics analysis methods [20], many entomologists use various ensemble research methods such as

genome, transcriptome, proteome, etc. to obtain a large number of molecular biological data, and use bioinformatics methods to analyze and mine the data. This paper attempts to reveal the genetic basis of the special physiological activities and behaviors of some insects from the point of view of molecular biology [21,22], expounding the molecular mechanism of co-evolution of insects and other organisms [23], showing a broad research prospect of insect molecular biology [24,25].

Coleoptera is the first order with the largest number of species and the widest distribution in the Insecta and even the animal kingdom. There are many kinds and the system is complex. However, there are also many species (mainly larvae) that harm crops and bring loss or inconvenience to human production and life [26]. Therefore, the research on Coleoptera insect genomics has important theoretical significance and application value. Nowadays, the genomic research of Coleoptera insects is mainly to analyze their special habits or behaviors at the genome level [27-30]. Genome size, also known as C-value or Constant-value, is an important
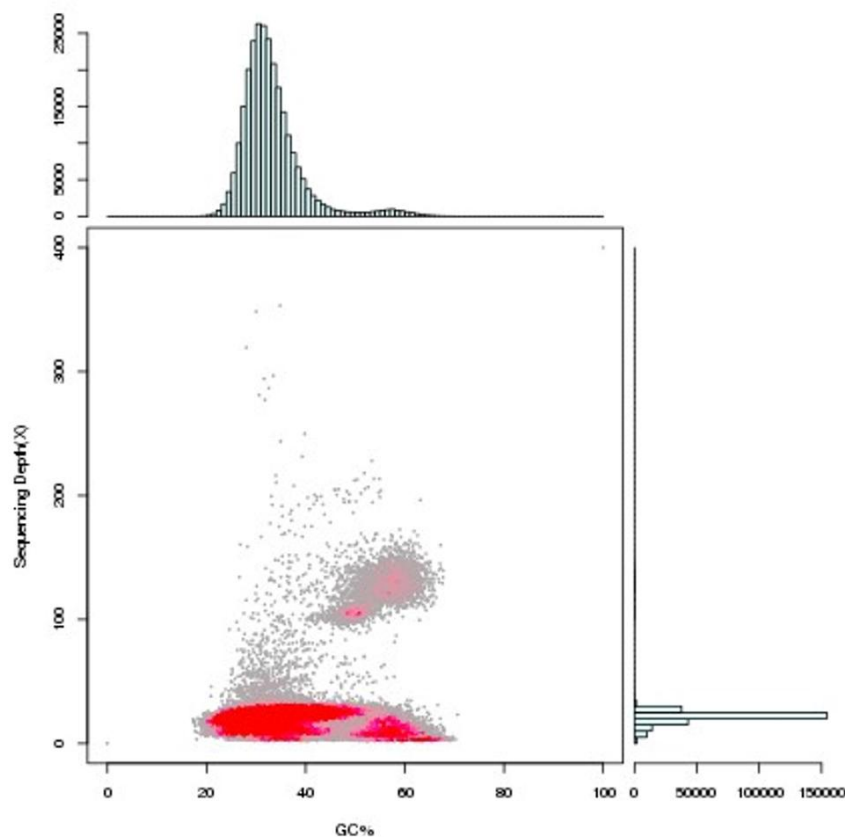

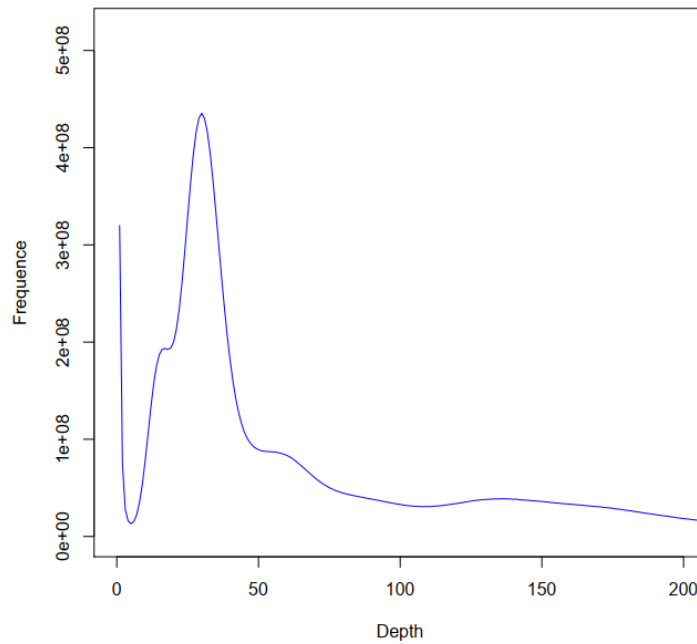**Fig. 2. GC content and average sequencing depth**

Fig. 3. 17 *K*-mer analysis for estimating the genome size of *Monochamus alternatus*

genetic feature of an organism. *M. alternatus* is known as the cancer of pine trees and is also a devastating disease, causing serious environmental and economic losses. Effective strategies are needed to stop or control the spreading of this disease. This study will provide a method to assemble the genome and may serve as key point to develop new control strategies for pine wilt disease, the results of the current study will play an important role in future whole-genome sequencing projects and provide an abundant resource for further functional studies, which will help us learn genetic regulatory mechanisms.

Table 2. Comparison of the mitogenome assembly data of *Monochamus alternatus* with that of 25 species of Coleoptera

Species	Genome size (Mb)	Number of protein-coding genes	GC content (%)	Heterozygosity (%)	Repeat (%)	References
Monochamus alternatus	871.09	n.a.	32.94	1.04	59.04	This study
Diabrotica virgifera	2418.07	28061	36.5	n.a.	n.a.	[15]
Photinus pyralis	471.51	32294	36.4	0.598	42.6	[31]
Listronotus bonariensis	1112.44	n.a.	31.3	0.18	70	[30]
Propylea japonica	851.23	18018	35.13	0.9	58.22	[32]
Marronus borbonicus	406.94	23278	35.9	0.2	29.2	[33]
Abscondita terminalis	499.65	20439	31.4	n.a.	n.a.	[34]
Lamprigera yunnana	1052.93	19438	34.1	n.a.	n.a.	[34]
Hycleus cichorii	99.17	13813	32.3	1.00	72.73	[35]

| Species	Genome size (Mb)	Number of protein-coding genes	GC content (%)	Heterozygosity (%)	Repeat (%)	References
Harmonia axyridis	466.692	n.a.	33.2	n.a.	n.a.	[36]
Hycleus phaleratus	90.87	13725	30.06	0.99	74.90	[35]
Aethina tumida	234.34	17463	30.00	n.a.	n.a.	[37]
Limonius californicus	1072.67	n.a.	35.00	0.21	n.a.	[38]
Oryctes borbonicus	406.20	8822	34.85	0.2	29.2	[33]
Nicrophorus vespilloides	195.27	18995	32.20	n.a.	n.a.	[22]
Asbolus verrucosus	249.61	n.a.	32.90	n.a.	n.a.	[39]
Protaetia brevitarsis	1143.92	n.a.	25.40	n.a.	72.29	[40]
Popillia japonica	531.53	n.a.	34.90	n.a.	n.a.	[41]
Anoplophora glabripennis	706.97	20632	33.40	n.a.	n.a.	[25]
Tenebrio molitor	280.78	n.a.	26.50	n.a.	n.a.	[42]
Agrilus planipennis	353.07	22159	36.00	n.a.	n.a.	[17]
Leptinotarsa decemlineata	641.99	19038	35.60	n.a.	n.a.	[43]
Callosobruchus maculatus	1007.82	31345	37.7	n.a.	65.00	[44]
Priacma serrata	12.08	n.a.	35.90	n.a.	n.a.	[45]
Dendroctonus ponderosae	257.09	16791	38.45	n.a.	n.a.	[33]
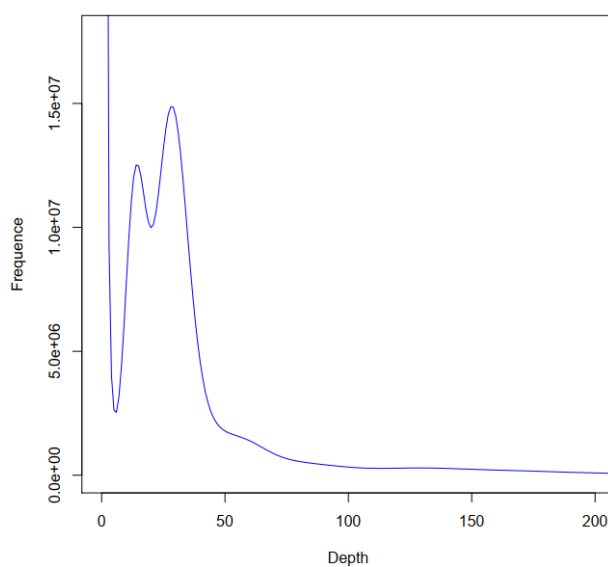
Fig. 4. 17 *K*-mer analysis for estimating the genome size of *Monochamus alternatus*
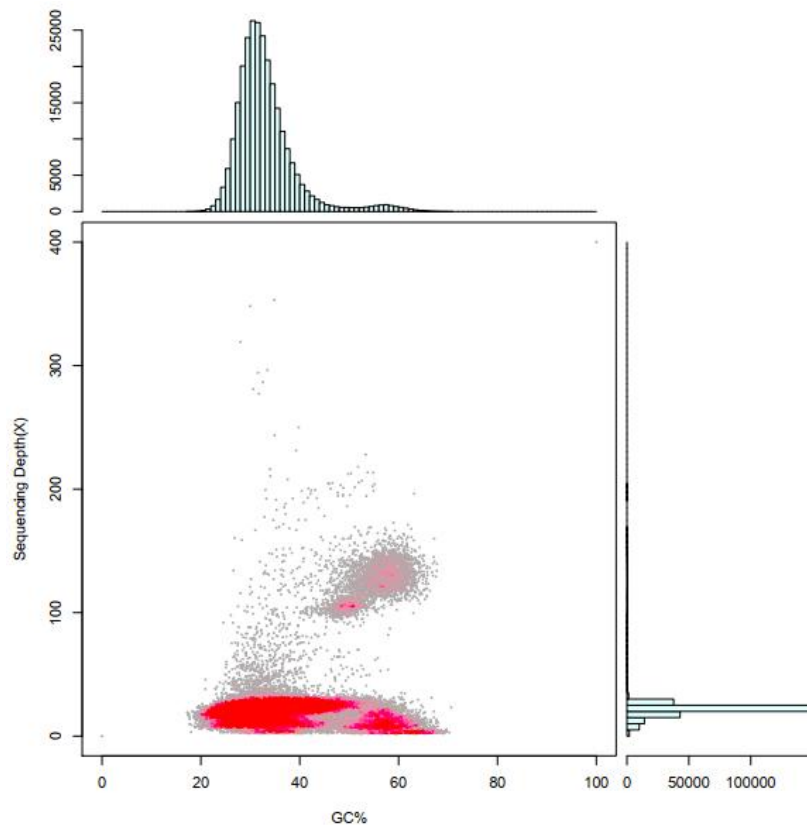
Fig. 5. GC content and average sequencing depth

This study lays the foundation of whole-genome sequencing in *M. alternatus*, and puts forward the platform for the next investigation on this particular insect. Today, more and more scholars studied the relationship between *M. alternatus* and pine wood nematode, because of its great harm to the pine wood. Our study increases the abundance of the Cerambycidae genome information and can assist in phylogenetic, molecular systematics and evolutionary studies of Cerambycidae. However, molecular data on species of Lamiinae are still scarce and more information is needed to fully explore the relationships within Lamiinae.

5. CONCLUSION

The sequencing quality of this study was good and within the normal range. The *K*-mer data indicated that there is a large proportion of repeats in the genome of *M. alternatus*. The genome GC content is moderate, which will not affect the accuracy and correct assembly of genome sequencing. The genome of *M. alternatus* is a complex genome with high repeat sequences.

ACKNOWLEDGEMENTS

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Nie R, Vogler AP, Yang XK, Lin MY. Higher-level phylogeny of longhorn beetles (Coleoptera: Chrysomeloidea) inferred from mitochondrial genomes. Syst. Entomol. 2020;46: 56–70.
2. Wang J, Dai XY, Xu XD, Zhang ZY, Yu DN, Storey KB, Zhang JY. The complete mitochondrial genomes of five longicorn beetles (Coleoptera: Cerambycidae) and phylogenetic relationships within Cerambycidae. PeerJ. 2019;7:e7633.

3. Santana DS, Marinoni L, Monné ML, Gómez-Zurita J. Molecular phylogenetic assessment of the tribal classification of Lamiinae (Coleoptera: Cerambycidae). Mol. Phylogenet. Evol. 2020;145:106736.

4. Min J, Kim HHI, Kwang YL, Yeon SH, Iksoo K. Complete mitochondrial genome of the whiter-spotted flower chafer, Protaetia brevitarsis (Coleoptera: Scarabaeidae). Mitochondrial DNA 2014; 25(3):177–178.

5. Wen Z, Yang YZ, Shi QF, Li JL, Zhi HL. The first complete mitochondrial genome of the Japanese beetle Popillia japonica (Coleoptera: Scarabaeidae) and its phylogenetic implications for the superfamily Scarabaeoidea. Int. J. Biol. Macromol. 2018;118:1406–1413.

6. Yang F and He DH. Effects of habitat fragmentation on biodiversity. Ann. Rev. Ecol. Evol. S. 2003;34(02):487-515.

7. Chen WB, Xiao DN, Li XZ. Classification, application, and creation of landscape indices J. Appl. Entomol. 2012;13(01):121-125.

8. Hu SJ, Ning T, Fu DY, Haack RA, Zhang Z, Chen DD, Ma XY, Ye H. Dispersal of the Japanese pine sawyer, *Monochamus alternatus* (Coleoptera: Cerambycidae), in mainland China as inferred from molecular data and associations to indices of human activity. PLoS ONE 2013;8(02):e57568.

9. Xu XM. Effects of environmental conditions on infection of Fusarium ear blight. Eur. J. Plant Pathol. 2013;109(07):683-389.

10. Haddad S, Shin S, Lemmon AR, Lemmon EM, Svacha P, Farrell B, Windsor D, Mckenna DD. Anchored hybrid enrichment provides new insights into the phylogeny and evolution of longhorned beetles (Cerambycidae). Syst. Entomol. 2018; 43:68–89.

11. Dai XY, Zhang H, Xu XD, Jia YY, Zhang JY, Yu DN, Cheng HY. The complete mitochondrial genome of Annamanum lunulatum (Coleoptera: Lamiinae) and its phylogeny. Mitochondrial DNA Part B. 2020;5:551–553.

12. Shi H, Zhang CQ, Shan L, Qi Q. First report of Pestalo-tiopsis microspora as a causal agent of black spot of pecan (Carya illinoensis) in China. Plant Dis. 2015; 9(09):1276.

13. Bai L, Tian CM, Hong CH, Kang FF, Chen JY, Song DW, Liu HG. The relationship between pine forest landscape patterns and pine wilt disease in Yichang of Hubei Province. Acta Ecologica Sinica. 2015; 35(24):8107-8116.

14. Zhou J, Zhao LL, Yu HY, Wang YH, Zhang W, Hu SN, Zou Z, Sun JH. Immune tolerance of vector beetle to its partner plant parasitic nematode modulated by its insect parasitic nematode. FASEB J. 2018;32:4862–4877.

15. Wu K, Camargo C, Fishilevich E, Narva KE, Chen X, Taylor CE, Siegfried BD. Distinct fitness costs associated with the knockdown of RNAi pathway genes in western Corn rootworm adults. PLoS ONE 2017;12(12):e0190208.

16. Sharkey CR, Fujimoto MS, Lord NP, Shin S, McKenna DD, Suvorov A, Martin GJ, Bybee SM. Overcoming the loss of blue sensitivity through opsin duplication in the largest animal group, beetles. Sci. Rep. 2016;7:8.

17. Lord NP, Plimpton RL, Sharkey CR, Suvorov A, Lelito JP, Willardson BM, Bybee SM. A cure for the blues: opsin duplication and subfunctionalization for short-wavelength sensitivity in jewel beetles (Coleoptera: Buprestidae). BMC Evol. Biol. 2016;16:107.

18. Nakamatsu Y and Tanaka T. The function of a trypsin－like enzyme in the saliva of Euplectrus separatae larvae. J. Insect Physiol. 2004;50(09): 847-854.

19. Shah K, Alqudah MA, Jarad F, Abdeljawad T. Semi-analytical study of Pine Wilt Disease model with convex rate under Caputo–Febrizio fractional order derivative. Chaos. Soliton. Fract. 2020;135:109754.

20. Wu S, Gao S, Wang S, Meng J, Wickham J, Luo S, Tan X, Yu H, Xiang Y, Hu S, et al. A reference genome of Bursaphelenchus mucronatus provides new resources for revealing its displacement by pinewood nematode. Genes. 2020;11: 570.

21. Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N, Donahue G, Yang P, Li Q, Li C, et al. Genomic comparison of the ants Camponotus floridanus and Harpegnathos saltator. Science 2010;329(5995): 1068-1071.

22. Cunningham CB, Ji L, Wiberg RA, Shelton J, McKinney EC, Parker DJ, Meagher RB, Benowitz KM, Roy-Zokan EM, Ritchie MG, et al. The Genome and methylome of a beetle with complex social Behavior, Nicrophorus vespilloides (Coleoptera: Silphidae). Genome Biol. Evol. 2015; 7(12):3383–3396.

23. Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, Li B, Cui F, Wei J, Ma C, et al. The locust genome provides insight into swarm formation and long-distance flight. Nat. Commun. 2014;5:2957-2965.

24. Ren BZ, and Tang LH. Progress in quarantine pests of Bursaphelenchus xylophilus. J. Jilin Agric. Univ. 2020;42: 8–13.

25. Lowe TM and Sean RE. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25(05):955–964.

26. Zhang JY, Luu BE, Yu DN, Zhang LP, Al-attar R, Storey KB. The complete mitochondrial genome of Dryophytes versicolor: Phylogenetic relationship among Hylidae and mitochondrial protein-coding gene expression in response to freezing and anoxia. Int. J. Biol. Macromol. 2019;132:461–469.

27. Zhan S, Merlin C, Boore JL, Reppert SM. The monarch butterfly genome yields insights into long-distance migration. Cell 2011;147(05):1171-1185.

28. You MS, Yue Z, He W, Yang X, Yang G, Xie M, Zhan D, Baxter SW, Vasseur L, Gurr GM, et al. A heterozygous moth genome provides insights into herbivory and detoxification. Nat. Genet. 2013; 45(02):220-225.

29. Zhan S, Zhang W, Niitepõld K, Hsu J, Haeger JF, Zalucki MP, Altizer S, de Roode JC, Reppert SM, Kronforst MR. The genetics of monarch butterfly migration and warning colouration. Nature 2014; 514(7522):317-31.

30. Harrop TWR, Le Lec MF, Jauregui R, Taylor SE, Inwood SN, van Stijn T, Henry H, Skelly J, Ganesh S, Ashby RL, et al. Genetic diversity in invasive populations of Argentine stem weevil associated with adaptation to biocontrol. Insects 2020; 11:441.

31. Fallon TR, Lower SE, Chang CH, Bessho-Uehara M, Martin GJ, Bewick AJ, Behringer M, Debat HJ, Wong I, Day JC, et al. Firefly genomes illuminate parallel origins of bioluminescence in beetles. Elife 2018;7:e36495.

32. Zhang L, Li S, Luo J, Du P, Wu L, Li Y, Zhu X, Wang L, Zhang S, Cui J. Chromosome-level genome assembly of the predator Propylea japonica to understand its tolerance to insecticides and high temperatures. Mol Ecol Resour. 2020;20: 292-307.

33. Meyer JM, Markov GV, Baskaran P, Herrmann M, Sommer RJ, Rödelsperger C. Draft genome of the scarab beetle Oryctes borbonicus on La Réunion Island. Genome Biol Evol. 2016;8:2093–2105.

34. Zhang R, He J, Dong Z, Liu G, Yin Y, Zhang X, Li Q, Ren Y, Yang Y, Liu W, et al. Genomic and experimental data provide new insights into luciferin biosynthesis and bioluminescence evolution in fireflies. Sci Rep. 2020;10: 15882.

35. Wu YM, Li J, Chen XS. Draft genomes of two blister beetles Hycleus cichorii and Hycleus phaleratus. Gigascience. 2018;7: 1-7.

36. Ando T, Matsuda T, Goto K, Hara K, Ito A, Hirata J, Yatomi J, Kajitani R, Okuno M, Yamaguchi K, et al. Repeated inversions within a pannier intron drive diversification of intraspecific colour patterns of ladybird beetles. Nat Commun. 2018;9:3843

37. Evans JD, McKenna D, Scully E, Cook SC, Dainat B, Egekwu N, Grubbs N, Lopez D, Lorenzen MD, Reyna SM, et al. Genome of the small hive beetle (Aethina tumida, Coleoptera: Nitidulidae), a worldwide parasite of social bee colonies, provides insights into detoxification and herbivory. Gigascience. 2018 12 01;7: giy138.

38. Andrews KR, Gerritsen A, Rashed A, Crowder DW, Rondon SI, van Herk WG, Vernon R, Wanner KW, Wilson CM, New DD, et al. Wireworm (Coleoptera: Elateridae) genomic analysis reveals putative cryptic species, population structure, and adaptation to pest control. Commun Biol. 2020 09 07;3:489.

39. Rider SD. The complete mitochondrial genome of the desert darkling beetle Asbolus verrucosus (Coleoptera, Tenebrionidae). Mitochondrial DNA A DNA Mapp Seq Anal. 2016;27(4):2447-9.

40. Kim MJ, Im HH, Lee KY, Han YS, Kim I. Complete mitochondrial genome of the whiter-spotted flower chafer, Protaetia brevitarsis (Coleoptera: Scarabaeidae). Mitochondrial DNA. 2014;25:177-8.

41. Yang W, Zhang Y, Feng S, Liu L, Li Z. The first complete mitochondrial genome of the Japanese beetle Popillia japonica (Coleoptera: Scarabaeidae) and its phylogenetic implications for the superfamily Scarabaeoidea. Int J Biol Macromol. 2018;118:1406-1413.

42. Liu LN, Wang CY. Complete mitochondrial genome of yellow meal worm (*Tenebrio molitor*). Dongwuxue Yanjiu. 2014;35: 537-45.

43. Crossley MS, Chen YH, Groves RL, Schoville SD. Landscape genomics of Colorado potato beetle provides evidence of polygenic adaptation to insecticides. Mol Ecol. 2017;26: 6284-6300.

44. Sayadi A, Martinez Barrio A, Immonen E, Dainat J, Berger D, Tellgren-Roth C, Nystedt B, Arnqvist G. The genomic footprint of sexual conflict. Nat Ecol Evol. 2019;3:1725-1730.

45. Niehuis O, Hartig G, Grath S, Pohl H, Lehmann J, Tafer H, Donath A, Krauss V, Eisenhardt C, Hertel J, Petersen M, Mayer C, Meusemann K, Peters RS, Stadler PF, Beutel RG, Bornberg BE, McKenna DD, Misof B. Genomic and morphological evidence converge to resolve the enigma of Strepsiptera. Curr Biol. 2012;22(14): 1309-13.