Draft genome sequence of the Algerian bee *Apis mellifera intermissa*

Nizar Jamal Haddad\textsuperscript{a,1}, Wahida Loucif-Ayad\textsuperscript{b,c}, Noureddine Adjlane\textsuperscript{d,⁎}, Deepi Saini\textsuperscript{e}, Rushiraj Manchiganti\textsuperscript{e}, Venkatesh Krishnamurthy\textsuperscript{e}, Banan AlShagoor\textsuperscript{a,1}, Ahmed Mahmud Batainha\textsuperscript{a,1}, Raja Mugasimangalam\textsuperscript{e}

\textsuperscript{a} Bee Research Department, National Center for Agricultural Research and Extension, P.O. Box 639, Bâja’ 19381, Jordan
\textsuperscript{b} Laboratory of Applied Animal Biology, Faculty of Science, University Badji-Mokhtar, Annaba, Algeria
\textsuperscript{c} Faculty of Medicine, University Badji-Mokhtar, Annaba, Algeria
\textsuperscript{d} Department of Biology, Faculty of Science, University M‘hamed Bougara of Boumerdes, Boumerdes, Algeria
\textsuperscript{e} Research and Development Unit, Genotypic Technology (P) Ltd.,#259, Apurva, 2nd Floor, 4th Cross, R.M.V. 2nd Stage, 80 Feet Road, Bangalore 560 094, India

**Abstract**

*Apis mellifera intermissa* is the native honeybee subspecies of Algeria. *A. m. intermissa* occurs in Tunisia, Algeria and Morocco, between the Atlas and the Mediterranean and Atlantic coasts. This bee is very important due to its high ability to adapt to great variations in climatic conditions and due to its preferable cleaning behavior. Here we report the draft genome sequence of this honey bee, its Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JSUV00000000. The 240-Mb genome is being annotated and analyzed. Comparison with the genome of other *Apis mellifera* sub-species promises to yield insights into the evolution of adaptations to high temperature and resistance to Varroa parasite infestation.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
before subjecting it to the assembly software. After the primary analysis of the assembly described below, based on contig sizes, genome representation and its functional elements, the output will be taken up for further analysis.

**Raw data QC**

The Illumina GAIIx paired end raw reads were quality checked using Genotypic Pvt. Ltd., proprietary tool SeqQC. Illumina library has uniform read length of 100.

**Raw data processing**

Illumina paired-end raw reads were processed by perl script for removal of adapters, low-quality bases, and b-blocks trimming towards 3′-end of the read. Coverage is calculated on the assumption of the 240 Mb genome size of *A. m. intermissa* as shown in Supplement.

**De-novo assembly of sequenced genome data**

Short reads paired-end read data, 85.5 million read pairs were assembled with SOAP denovo assembler [14]. Various hash-lengths were tested for denovo assembly, ranging from 45 to 65 km. The result shows that 59 hash-length has assembled the best assembly in terms of total contig length, number of contigs generated, and N50 value. The denovo assembly generated 2330251 contigs. SOAP denovo was used with default parameters. Scaffolding was completed using the software SSPACE 2.0 scaffolder using paired-end data [15]. Scaffolding resulted in 522,976 scaffolds. Although SOAP denovo uses paired reads to resolve repeats and polymorphism, it does not perform “gapped” scaffolding: if there is a coverage gap between a read pair, this pair is not utilized. To leverage such pairs and to make better use of long insert libraries, scaffolding is carried out. SSPACE scaffolder is used with these parameters: minimum number of overlapping bases with the seed is 45, minimum overlap required between contigs to merge adjacent contigs in a scaffold is 50, minimum read pairs to compute scaffold is 5, contig extension is switched on, minimum number of reads needed to call a base during an extension is 20, and maximum number of allowed gaps during mapping with Bowtie Langmead et al. [16] is 1. Scaffolds do consist of uncalled bases (Ns). To fill these inter-scaffold Ns with nucleotides, gap closure is done using Gap Closer tool later. With the help of SSPACE scaffolding using Illumina library, the assembly showed significant improvement in N50 value and significant reduction in number of contigs. The draft genome (scaffolds) resulted 522,976 sequences which gives a total genome length of 243 Mb as shown in Supplement.

The draft genome information reported here provides opportunity for further research into the mechanism involved in adaptations that allow the organisms to thrive in high temperature environments and how these bees have good immunity and are resistant to infestation of Varroa parasite.

**Summary**

*Apis m. intermissa* is the native honeybee subspecies of Algeria. *A. m. intermissa* occurs in Tunisia, Algeria and Morocco, between the Atlas and the Mediterranean and Atlantic coasts [12], in an area of more than 2500 km long. *A. m. intermissa* indicates the position through this bee races between tropical Africa and European breeds [17]. This bee is very aggressive, nervous and produces many broods with many queen cells [6]. *A. m. intermissa* is prone to swarming, shows a defensive behavior and an abundant use of propolis [12]. Here we report the draft genome sequence of this honey bee. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JSUV00000000 (http://www.ncbi.nlm.nih.gov/nuccore?JSUV00000000). The version described in this paper is version JSUV01000000. The 240-Mb genome is being annotated and analyzed. Comparison with the genome of other *Apis m. intermissa* sub-species promise to yield insights into the evolution of adaptations to high temperature and resistance to Varroa parasite infestation.

**Competing interests**

The authors have declared that no competing interest exists.

**Acknowledgments**

We would like to thank the Algerian beekeepers Hamzouei M. and Diffalaf H (Association for the Development of beekeeping in the Mitidja: ADAM), for their intensive collaboration in honeybee sampling. This study is part of the project entitled “Evaluation and Development of *Apis mellifera syriaca* with Useful Beekeeping Traits for the Middle East Region.” funded by USAID grant number: TA-MOU-09-M29-075.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2015.01.011.

**References**

[1] N. Adjlane, S. Kechih, N. Haddad, Comparative study between techniques for the diagnosis of American foulbrood (*Paenibacillus larvae*) in honeybee colony, J. Anim. Vet. Adv. 13 (16) (2014) 970–973.

[2] N. Adjlane, N. Haddad, The first data on hygienic behavior of *Apis mellifera intermissa* in Algeria, J. Biol. Earth Sci. 4 (1) (2014) 1–5.

[3] N. Adjlane, T. El-Ounass, N. Haddad, Population Dynamics of Varroa destructor in Colonies of *Apis mellifera intermissa* in Algeria. 10th COLOSS Conference, 6-8th September 2014, Murcia/Spain, 2014.

[4] C. Barou, A. Tahar, M. Baylac, Forewinding shape variation in Algerian honey bee populations of *Apis mellifera intermissa* (Buttel-Reepen, 1906) (*Hymenoptera: Apidae*); a landmark-based geometric morphometrics analysis. Afr. Entomol. 19 (2011) 11–22.

[5] G. Boecking, W. Ritter, Grooming and hygienic behaviour of *Apis mellifera intermissa* in Tunisia against *Varroa jacobsoni*. J. Apic. Res. 32 (1993) 127–134.

[6] H. Clément, E. Bruneau, J.M. Barançon, P. Bonnafé, R. Domérgue, G. Fert, Y. Le Conte, J. Ratia, C. Reeb, B. Vassière, Le traité Rustica de l’apiculture. Traité Rustica, Paris, 2002. 528.

[7] C. Gadbin, J.M. Cornuet, J. Fresnaye, Approche biométrique de la variété locale d’*Apis mellifera* I dans le sud tchadien. Apidologie 10 (1979) 137–148.

[8] J.A. Keřus, Honey bee hygienic behavior; France, Tunisia and Chile. Apidologie 26 (1995) 24–26.

[9] W. Loucif-Ayad, A. Cherfouri, M. Alburaki, L. Garnery, Genetic assessment of Algerian honeybee populations by microsatellite markers. Apidologie (2014) http://dx.doi.org/10.1051/aps/2013052.

[10] W. Loucif-Ayad, M. Aouch, H. Legout, M. Alburaki, L. Garnery, Genetic assessment of Algerian honeybee populations by microsatellite markers. Apidologie (2014) http://dx.doi.org/10.1051/aps/2013052.

[11] F. Rutten, Biogeography and Taxonomy of Honey Bees. Springer, Berlin, 1988.

[12] T. Shalbi, S. Fuchs, R.F.A. Moritz, Morphological studies of honeybees (*Apis mellifera*) from Libya. Apidologie 40 (2) (2009) 97–105.

[13] R. Luo, B. Liu, Y. Xin, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu, J. Tang, G. Wu, H. Zhang, Y. Shi, Y. Liu, C. Yu, B. Wang, Y. Lu, C. Han, D.W. Cheung, S.M. Yu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang, T.W. Lam, J. Wang, SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Giga Sci. 1 (1) (2012) 16.

[14] M. Boetzer, W. Piwowar, SPACeLongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinf. 15 (2014) 211. http://dx.doi.org/10.1186/1471-2105-15-211.

[15] B. Langmead, C. Trapnell, M. Pop, S.J. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10 (2009) R25.

[16] Ch. Peyvet, L’espèce *Apis mellifera* : les grandes races géographiques. Bull. Tech. Apic. 21 (1994) 129–138.