



HAL
open science

Tenebrio molitor genomic structure among available populations.

Evangelia Eleftheriou, Benoît Vacherie, Karine Labadie, C.G. Athanassiou, Thierry Rigaud, Yannick Moret, Thomas Lefebvre, Mohammed-Amin Madoui

► **To cite this version:**

Evangelia Eleftheriou, Benoît Vacherie, Karine Labadie, C.G. Athanassiou, Thierry Rigaud, et al.. Tenebrio molitor genomic structure among available populations.. WCGALP 2022 12th World Congress on Genetics Applied to Livestock Production, Jul 2022, Rotterdam, Netherlands. pp.2549-2551, <10.3920/978-90-8686-940-4_616>. <hal-04196166>

HAL Id: hal-04196166

<https://hal.science/hal-04196166v1>

Submitted on 5 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

616. *Tenebrio molitor* genomic structure among available populations

E. Eleftheriou¹, B. Vacherie², K. Labadie², C.G. Athanassiou³, T. Rigaud⁴, Y. Moret⁴, T. Lefebvre⁵ and MA. Madoui^{1,6*}

¹Génomique Métabolique, Genoscope, Institut François Jacob, Commissariat à l'Energie Atomique (CEA), CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France; ²Genoscope, Institut de biologie François Jacob, CEA, Université Paris-Saclay, 91057 Evry, France; ³Department of Agriculture Crop Production and Rural Environment, School of Agricultural Sciences, University of Thessaly, Fytokou Str., 38446 Volos, Greece; ⁴UMR CNRS 6282 BioGéoSciences, Équipe Écologie Évolutive, Université Bourgogne Franche Comté, 21000 Dijon, France; ⁵Ynsect, 91000 Evry, France; ⁶Service d'Etude des Prions et des Infections Atypiques (SEPIA), Institut François Jacob, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Université Paris Saclay, Fontenay-aux-Roses, France; mohammedamin.madoui@cea.fr

Abstract

Tenebrio molitor is now a promising species for insect farming to produce high protein products with low environmental impact. However, the molecular diversity and the genetic structure in available *T. molitor* populations are understudied. We analysed the genomic structure of 18 populations from 12 countries using Illumina pool-seq data. More than 5 million SNPs were selected to compute allele frequencies, then multivariate and pairwise- F_{ST} analyses were used to estimate their relative genetic distances. The results showed two very distant populations, one wild population sampled in Serbia and one from Germany, both presenting very distinct genomic structure. The populations from USA, Spain and Holland presented moderate genomic differentiation and 13 populations were closely-related. This pool-seq analysis highlighted a lack of molecular diversity between most of the sampled populations and will allow the identification of loci under selection. Current phenotyping and genotyping efforts will help guiding future breeding programs.

Introduction

By 2050, the global population will reach approximately nine billion people, increasing the demand for animal protein by 76% (Alexandratos and Bruinsma, 2012). This future demand starts to question the sustainability of our conventional production systems. Nowadays, insect production is considered as a serious alternative for food and feed production for many reasons. Essentially for the nutritional composition of edible insects (Nowak *et al.*, 2016) and for the low environmental impact of its production compared to other conventional livestock production systems (van Huis, 2013).

The yellow mealworm beetle *Tenebrio molitor* has been described as a good candidate for animal and even human nutrition (Cortes Ortiz *et al.*, 2016). Thus, genomic efforts and industrial scale production of *T. molitor* progressed during the last decade (Eleftheriou *et al.*, 2021). Despite promising results on *T. molitor* production, several aspects concerning the *T. molitor* biology remain challenging to sustain its production (Morales-Ramos *et al.*, 2019). The need to have access to the molecular and phenotypic diversity of *T. molitor* is a first step before further genetic improvement. Indeed, industrial farming of *T. molitor* will require monitoring the molecular diversity in order to better anticipate inbreeding depression that may cause decrease of biomass production. Here, we used Illumina pool-seq on 18 *T. molitor* populations and sequenced 720 individuals to identify their molecular diversity and genomic structure.

Materials & methods

Biological material. A total of 18 populations were provided by academic institutions and private companies. University of Burgundy (Dijon, France) provided three populations, one called *Dijon*, one called *Serbia* that corresponds to a wild population sampled in a disaffected mill in Serbia and one called

Holland that corresponds to a population initially provided by a private Dutch company. University of Thessaly (Volos, Greece) provided seven populations, two populations from Italy called *Italy_1* and *Italy_2*, one population from Greece called *Greece*, one population from Spain called *Spain*, one population from Germany called *Germany*, one population from the United States of America called *USA* and one population from Turkey called *Turkey*. Mystik Farms (Cap Santé, Québec, Canada) provided one population called *Canada*. Papek s.r.o. (Jaroměřice nad Rokytnou, Czech Republic) provided one population called *Czech*. Two populations called *Pro_+* and *Pro_-* were provided by Pronutrix (Harelbeke, Belgium). One population called *Belgium* was provided by Nusect (Deerlijk, Belgium). The Ynsect company (Evry, France) provided one population called *AAF_Dole*. The Kinsect company (France) provided one population called *Kynsect_HOL*. The Star Food company (Barneveld, Netherlands) provided one population called *Starfood_HOL*.

Pool-seq sequencing. *Tenebrio molitor* larvae were first ground individually using ZR BashingBead Lysis Tubes (2 mm) (ZymoResearch, CA, USA) and subsequently DNA extraction was performed using the Quick-DNA 96 Plus Kit (ZymoResearch). The extracted DNA was normalized and 40 individuals from the same population were pooled equimolarly. For each pool, DNA libraries were prepared using the NEBNext® Ultra II DNA Library prep kit for Illumina (New England Biolabs, MA, USA) and sequenced on an Illumina NovaSeq6000 instrument (Illumina, San Diego, CA, USA) using 150 base-length read chemistry in paired-end mode to reach around 20x of depth of coverage by individual.

Population genomic analysis. Raw reads were cleaned and aligned against the *T. molitor* reference genome using *bwa mem* (Li and Durbin, 2009). Reads duplicates were removed and SNPs were called using samtools/bcftools. For each SNP, the alternative allele frequency was calculated, to get rid of sequencing errors, only SNPs with a minimum frequency of 5% in at least one population were kept. The relatedness between populations was first observed by principal component analysis (PCA) using the allele frequencies as input data. Then, the genomic distance between each population was estimated by computing the pairwise- F_{ST} , first we computed the F_{ST} for each allele as follow: $FST = var(p) / (mean(p) \times (1 - mean(p)))$, where p is a vector of alternative allele frequencies of two populations. Second, we used the median F_{ST} over all alleles to compute the final pairwise- F_{ST} .

Results & discussion

We obtained more than 3 million SNPs from each pool of individuals. The PCA analysis based on the allele frequencies of 5,285,733 SNPs showed that the two first PCA axes (explaining only 36% of the total variance) are strongly linked to *Serbia* and *Germany* populations while the other 16 populations are clustered into one single group located close to the origins of the axes (Figure 1.A). This suggests that *Serbia* and *Germany* have very different genomic structures while the 16 other populations may have closer genomic structures.

The pairwise- F_{ST} analysis showed that *Serbia* presented F_{ST} values ranging from 0.09 to 0.34 with the other populations, the larger value corresponding to the F_{ST} with *Germany* (Figure 1.B). *Germany* presented F_{ST} values ranging from 0.25 to 0.34. *Kynsect_HOL*, *USA* and *Spain* pairwise- F_{ST} values between 0.03 and 0.12. The last 13 populations presented low genomic differentiation between them with F_{ST} value ranging from 0 to 0.06. The F_{ST} analysis confirmed the PCA results by highlighting *Serbia* and *Germany* as two populations with high genomic differentiation compared to the 16 other populations but also validates the presence of different genomic structures between these two populations. But the F_{ST} -based analysis showed that a moderate differentiation occurred in *Spain*, *USA* and *Kynsect_HOL* and a low differentiation between the 13 other populations whose eight were sampled from private breeders. The recent interest in mealworm production by private companies may have led them to supply recently on the same or closely-related suppliers explaining possibly the observed low differentiation between their populations. On the other hand, evolutionary forces such as genetic drift and/or selection may also have occurred in the five more differentiated populations especially for *Serbia* that was sampled in a disaffected mill.

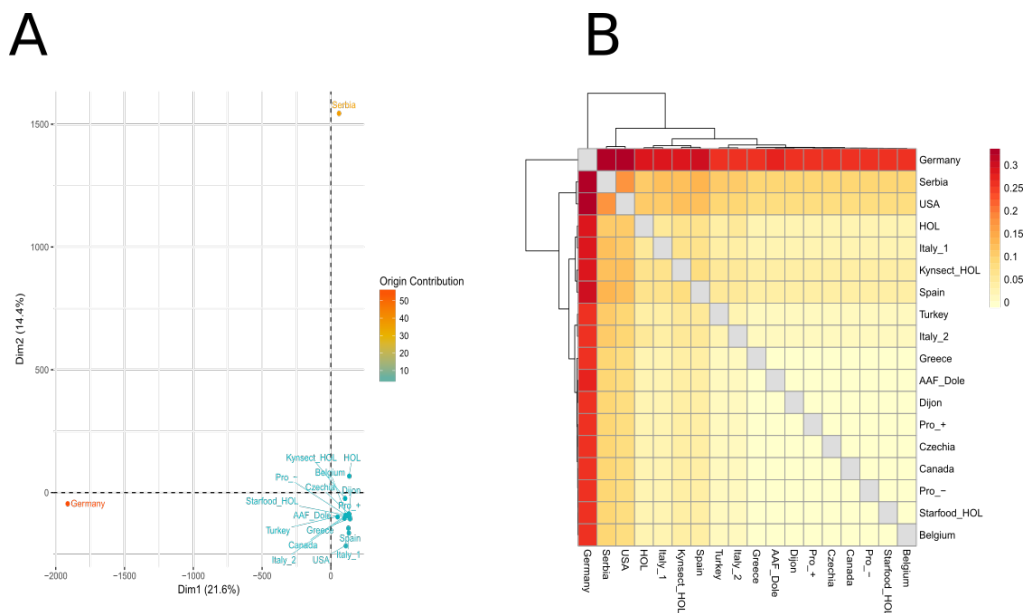


Figure 1. Genomic structure in 18 *T. molitor* populations. (A) PCA based on allele frequencies. (B) Pairwise- F_{ST} distance matrix.

Conclusions

From the pool-seq analysis presented in this study, we identified two populations having strong differentiation compared to the other 16 populations. We may consider that our results highlighted that the majority of the sampled populations presents a low molecular diversity between them. Current phenotyping of complex traits such as growth rate, fertility and immunity in available populations as well as their genotyping will help the selection of the populations having the most interesting performances for industrial production and will guide future breeding programs. Finally, to better understand molecular evolution of *T. molitor*, further analysis will be performed from the data generated in this study to identify genes under natural selection and the corresponding biological processes.

References

- Alexandratos, N., and Bruinsma, J. (2012). Food and Agriculture Organisation (FAO). World agriculture towards 2030/2050: the 2012 revision. Available at: <https://www.fao.org/3/ap106e/ap106e.pdf>
- Cortes Ortiz, J.A., Ruiz, A.T., Morales-Ramos, J. A. Thomas, M., Rojas, M. G., et al. (2016). Insect Mass Production Technologies. In: Dossey, A.T., Morales-Ramos, Juan A., Rojas, M. Guadalupe. (eds) Insects as Sustainable Food Ingredients. Academic Press, San Diego, USA, pp. 153–201. <https://doi.org/10.1016/B978-0-12-802856-8.00006-5>
- Eleftheriou, E., Aury, J., Vacherie, B., Istace, B., Belsler, C., et al. (2021). Open Res. Eur. 1. <https://doi.org/10.12688/openreseurope.13987.1>
- Li, H., and Durbin, R. (2009). Bioinformatics 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Morales-Ramos, J.A., Kelstrup, H.C., Rojas, M.G., and Emery, V. (2019). J. Insect Sci. 19(2):4. <https://doi.org/10.1093/jisesa/iey110>
- Nowak, V., Persijn, D., Rittenschober, D., and Charrondiere, U.R. (2016). Food Chem. 193:39–46. <https://doi.org/10.1016/j.foodchem.2014.10.114>
- Van Huis, A. (2013). Annu. Rev. Entomol. 58:563–583. <https://doi.org/10.1146/annurev-ento-120811-153704>