Automatic determination of landmark coordinates for honeybee forewing venation using a new MATLAB-based tool

E. Ángel-Beamonte\textsuperscript{a}, P. Martín-Ramos\textsuperscript{b}, P. Santolaria\textsuperscript{a}, E. Sales\textsuperscript{a}, J. Abizanda\textsuperscript{a} and J.L. Yániz\textsuperscript{a*}

\textsuperscript{a}TECNOGAM research group, Environmental Sciences Institute (IUCA), Department of Animal Production and Food Sciences, University of Zaragoza, Huesca, Spain;
\textsuperscript{b}TECNOGAM research group, Environmental Sciences Institute (IUCA), Department of Agricultural and Natural Environment Sciences, University of Zaragoza, Huesca, Spain.

*Corresponding author: Jesús L. Yániz, Department of Animal Production and Food Sciences, EPSH, University of Zaragoza, Ctra. Cuarte S/N 22071 Huesca, Spain. E-mail: jyaniz@unizar.es; Tel.: +34 (974) 239312; Fax: +34 (974) 239302.
In this study, a novel and free-to-use MATLAB-based tool (WingMarks) is presented, aimed at facilitating the geometric morphometric analysis of forewing venation through an automatic recognition of vein junctions. Firstly, the ability of the new software to analyze wing images obtained by different optical equipment was evaluated. Even when a low-end USB-microscope was used for image collection, it reached 100% precision in the automatic detection of the landmarks for 74% of the samples, and most of failures corresponded to a single point and were easily corrected. The measurement error of WingMarks software was studied through repeated analysis of the same wing image, evincing that landmark determination was highly repeatable, even higher than that of widely used tpsDig manual software. In addition, a field test with 720 specimens from three subspecies (A. m. iberiensis, A. m. ligustica, and A. m. carnica) and from Buckfast hybrid, collected from 90 different colonies, was conducted. In conjunction with MorphoJ, the coordinates of the vein junctions extracted by the program led to an accurate classification of the specimens, confirmed by canonical variate analysis methods. WingMarks software can thus be deemed as a versatile, precise and accurate tool for the automatic recognition of Apis mellifera wing vein junctions, facilitating the identification of bee genetic diversity using geometric morphometrics. The program is available under Creative Commons license and can be easily adapted to different insect species.

Keywords: Apis mellifera; wing geometric morphometrics; landmarks; MATLAB; subspecies discrimination
Determinación automática de las coordenadas de los nodos de la venación alar de la abeja melífera utilizando una nueva herramienta basada en MATLAB

En este estudio se presenta una herramienta novedosa y de uso libre basada en MATLAB (WingMarks), cuyo objetivo es facilitar el análisis de la morfometría geométrica de la venación de las alas anteriores mediante un reconocimiento automático de las uniones venosas. En primer lugar, se evaluó la capacidad del nuevo programa para analizar imágenes de ala obtenidas por diferentes equipos ópticos. Incluso cuando se utilizó un microscopio USB de gama baja para la recogida de imágenes, se alcanzó una precisión del 100% en la detección automática de las intersecciones en el 74% de las muestras y la mayoría de los fallos correspondieron a un solo punto y fueron fáciles de corregir. El error de medición del programa WingMarks se estudió a través del análisis repetido de la misma imagen de ala, demostrando una alta repetibilidad, incluso mayor que la del programa manual tpsDig, ampliamente utilizado. Además, se realizó una prueba de campo con 720 abejas de tres subespecies (A. m. iberiensis, A. m. ligustica, y A. m. carnica) y del híbrido Buckfast, recogidas de 90 colonias diferentes. Las coordenadas de las uniones venosas extraídas por el programa condujeron, mediante el empleo de MorphoJ, a una clasificación precisa de las abejas, confirmada por métodos de análisis canónico de varianza. Por lo tanto, el programa WingMarks puede ser considerado como una herramienta versátil, precisa y exacta para el reconocimiento automático de las intersecciones venosas del ala en Apis mellifera, facilitando la identificación de la diversidad genética de las abejas utilizando la morfometría geométrica. El programa está disponible bajo licencia Creative Commons y puede ser fácilmente adaptado a diferentes especies de insectos.
Introduction

The genetic diversity of bees is essential, among other things, to allow better adaptation to regionally varying factors of climate and vegetation. However, intensive selection has enabled the production of improved lines whose mass importations in many countries have severely reduced the population of bees adapted to the local environment. In Europe and other parts of the world, there is an increasing trend towards uniformity of honey bee populations due to economically-driven processes (De la Rúa, Jaffé, Dall'Olio, Muñoz, & Serrano, 2009; Meixner et al., 2013).

Any study of genetic diversity must have tools that allow a reliable identification of the possible genetic variants to be obtained. Molecular genetics allows precise characterization of genetic diversity. However, its use requires highly specialized facilities, equipment and expertise, and it is time and resources consuming. A high degree of consistency between the geometric morphometrics of wing venation and molecular information has been demonstrated by Miguel et al. (2011). It has been recently determined that wing geometric morphometrics and microsatellite analysis provide similar discrimination of honey bee subspecies (Oleksa & Tofilski, 2015).

Therefore, wing venation morphometrics is particularly suitable to track phylogenetic relationships between subspecies. Its determination is economical and simple and can be done with little equipment. The major limitation of this technique is the need to precisely identify the venous wing intersections to obtain their coordinates.

In the geometric morphometrics of wing venation, a frequently used program is tpsDig (Rohlf, 2001), which requires the manual plotting of all vein intersections with a mouse-click. The main advantage of this software is its flexibility, but it requires a lot of human interaction that is prone to errors and reproducibility problems. Tofilski (2004)
developed the DrawWing open software, able to automatically determine coordinates of eighteen wing landmarks, which was claimed to reduce the analysis time and to improve precision in the identification of vein junctions. Nonetheless, in most articles on this subject, tpsDig software has remained the most popular choice (Charistos, Hatjina, Bouga, Mladenovic, & Maistros, 2014; Dolati, Rafie, & Khaledro, 2013; Francoy, Grassi, Imperatriz-Fonseca, de Jesús May-Itzá, & Quezada-Euán, 2011; Francoy et al., 2008; Francoy et al., 2009; Rasic, Mladenovic, & Stanisavljevic, 2015). ApiClass website (Baylac et al., 2008) has also been used in some studies (Barour, Tahar, & Baylac, 2011), but-as noted by Nawrocka, Kandemir, Fuchs, and Tofilski (2017)- it covers only a small fraction of honey bee subspecies, it does not allow the user to edit the coordinates in case of error and it does not return coordinate values (such functionality was available in the past, but has been deactivated). Other alternatives, such as ABIS (Francoy et al., 2008; Steinhage, Arbuckle, Schröder, Cremers, & D., 2001) are not available to the general public.

The aim of this study was to develop a new software based on MATLAB for the automatic recognition of wing vein junctions in *Apis mellifera* and to assess its precision and ability to discriminate between different genetic variants when combined with other freely available program packages for doing geometric morphometrics.

**Materials and methods**

**Sampling and processing**

Honey bee workers were collected during the spring of 2016 from commercial apiaries in Northeastern Spain. At least eight bees per colony were collected from the brood area. Bees were stored in ethanol (CAS No. 64-17-5, Sigma Aldrich, BioUltra, ≥99.8%) in a freezer at −20 °C until morphological examination was conducted. The right
forewing of each individual was cut close to its base and then transferred to 70:30, 50:50, and 20:80 (v/v) ethanol:water solutions for gradual hydration, and finally to distilled water. The wing was then carefully mounted between a microscope slide and a coverslip.

**Image acquisition**

Digital images of the wings were recorded by means of a setup composed of a microscope (DM4500B, Leica, Wetzlar, Germany) with a PLAN Apochromatic 1.25× CORR objective (Leica), and photographed with a Canon EOS 600D digital camera (Canon Inc., Tokyo, Japan). The camera was computer-controlled using EOS utility software (Canon Inc., Tokyo, Japan). In the first part of the study, digital images of the wings were also obtained using a simple USB-microscope (Handheld Digital Pro 44308, Celestron, Torrance, USA), adapted to a (bottom) illumination stage plate, where the slides with the wings were placed, and controlled with a computer using Celestron utility software.

**WingMarks software**

The geometric morphometric analysis of the images was conducted using a standalone executable file generated with MATLAB’s Compiler Toolbox on a desktop computer (Intel® Core™ I7-6700K CPU @ 4.00 GHz, 32.0 GB RAM) running Windows 10® x64). WingMarks software package was written in MATLAB® (The Mathworks Inc., Natick, MA, USA) v.8.6 (R2015b) using MATLAB’s proprietary language. A royalty-free standalone application was generated with MATLAB Compiler. The Windows® executable file (64-bit), distributed under Creative Commons (CC BY-NC-SA 4.0) license, can be downloaded free of charge from [http://iuca.unizar.es/wingmarks-software](http://iuca.unizar.es/wingmarks-software). The source code (which can be compiled in Linux® and MacOS® platforms,
with which MATLAB is also compatible) is also available upon request. A video tutorial is provided in the same webpage.

**Image analysis procedure**

For every forewing image, the coordinates of 19 landmarks located at vein intersections (Figure 1a), were automatically determined using WingMarks software. The position and numbering of the landmarks was the same as in Smith, Crespi, and Bookstein (1997). For processing purposes, images were adjusted to an 800 pixels width and converted into greyscale (Figure 1b).

Prior to their analysis, the right forewing of one of the specimens was loaded into the graphical environment in order to use it as a template for the subspecies under study (“Create new template” menu option). To correct image tilting, two points corresponding to the ends of the radial cell (12 and 15 in Figure 1a) were manually plotted. The subsequent spatial transformation based on these two points ensured that the 19 nodes of the wing venation always displayed a uniform and recognizable configuration. The coordinates of 19 vein intersections were then manually plotted and were stored (in a .mat file) for their subsequent use as a pattern for the automatic detection of the other wing samples.

The global image threshold to minimize the intraclass variance of the black and white pixels, calculated with Otsu's method (Otsu, 1975), was used for the image binarization (i.e., conversion to black and white). Morphological operations were conducted on the binary image to remove selected foreground pixels (thinning), obtaining 1-pixel thick lines. By reversing colors and removing spurious branches (using a median filter), a thinned image of the forewing venation with well-defined intersections was obtained (Figure 1c).
For the remaining wing images (“Gather image landmarks” menu option), only
the tilting correction had to be manually made. By using a spatial convolution of the
surroundings of the 19 retrieved points from the template with the new corrected image,
the corresponding landmarks in the new image were automatically detected (Figure 1d).
These landmarks and the neighboring regions were displayed to the user, who could
verify and, if necessary, correct their coordinates. Finally, coordinates were stored both
in a .mat file and in a .tps identified with the sample’s name (“Close sample & format
data” menu option), for their ulterior utilization in wing indexes calculation routines or
in external applications such as MorphoJ (Klingenberg, 2011).

The program also includes an option to calculate the cubital index, discoidal
shift angle and Hantel index (“Calculate three classical indexes” menu option),
exporting the results –not the coordinates– to a .ps file (“Generate results.ps file”); and
an option to delete the measurements done on certain image (“Delete measurements”),
which can be useful, for instance, to remove outliers detected in the classical indexes
results.

**Experimental design**

**Experiment 1**

The first experiment was designed to evaluate the reliability of WingMarks software. Its
ability to analyze wing images obtained by different optical equipment (Leica and USB-
Celestron microscopes) was checked. One hundred wings were photographed using
both microscopes and the accuracy in the vein intersections recognition was assessed
using both WingMarks software and ApiClass website (http://apiclass.mnhn.fr/). As
explained above, ApiClass website does not return coordinate values but allows the user
to visually check the automatic recognition of vein intersections, thus allowing a
comparison with the results from WingMarks software.

Moreover, the measurement error (i.e., variability in the coordinate
determination for a given wing) of WingMarks software in comparison with tpsDig was
studied through repeated analysis of the same wing image obtained using the Leica
microscope (3 wings, 30 repetitions/wing). The effect of rotation and displacement of
the wing in the image on the measurement error using WingMarks was also studied
using Leica and Celestron microscopes (3 wings, 30 repetitions rotated and placed
differently/wing and microscope). The variability of each coordinate was calculated
using coefficients of variation (CVs). Within-wing CVs were expressed as the mean of
individual values.

Experiment 2
This experiment was performed to compare the geometric morphometrics of different
subspecies/strains of bees using WingMarks software. A total of 90 colonies from 18
populations of *A. mellifera* were sampled, broken down as follows: subspecies *A. m.
iberiensis* (10), *A. m. ligustica* (2), *A. m. carnica* (2), and Buckfast hybrid (4). In total,
720 bees from 90 colonies were analyzed.

The landmark coordinates obtained from WingMarks software were
subsequently processed in MorphoJ package (Klingenberg, 2011). Alignment was
performed using Procrustes fit (Procrustes superimposition). Canonical variate analysis
(CVA), using subspecies/strain as the group variable, was performed after calculating
the average of all bees from each colony.
Results

Experiment 1

WingMarks software allowed the analysis of 100% of the wings using both microscopes. Using the Leica microscope-images, the precision of the automatic detection of the landmarks was 100% in 90% of the cases, while the detection failed in a single point in 9% of the samples and in two points in 1% of the samples. When the detection failed, the automatic landmark was close to the actual vein intersection, so a small shift was enough to correct it. The same Leica-images processed in ApiClass website allowed the analysis of 92% of the wings (those in which the program recognized 100% of the landmarks), while it failed in 4% of the samples in a single point and in 4% of the samples in two points. As ApiClass does not allow to correct landmarks, 8% of the wings processed using the Leica microscope had to be discarded.

The analysis of the wing images captured with the simple Celestron USB-microscope with WingMarks software revealed that the precision of the automatic detection of the landmarks was 100% in 87% of the cases, while the detection failed in a single point in 12% of the samples, and in two points in 1% of the samples. The same Celestron-images processed in the ApiClass website allowed to analyze 75% of the wings (those in which the program recognized 100% of the landmarks), while it failed in 15% of the samples in a single point and in 10% of the samples in two points. Since landmarks could not be corrected, 25% of the images of the wings collected with the Celestron microscope had to be discarded.

Landmark determination using WingMarks software was highly repeatable, with even lower coefficients of variation than those of tpsDig: within-wing CVs ranged from 0.06 to 0.12% for WingMarks and from 0.23 to 0.28% for tpsDig. Rotation and displacement of the wing in the image had little impact on WingMarks measurement.
error, as confirmed using both microscopes: within-wing CVs ranged from XXX to XXX% for the Leica microscope and from XXX to XXX% for the Celestron microscope, respectively.

Experiment 2
According to the results of colony CVA, the first, second and third axes explained 68.19%, 19.59% and 12.22% of the total variation, respectively. The scatter plot of colonies demonstrated high discrimination results between subspecies/strains (Figure 2). Cross-validation tests based on CVA correctly classified 100% of the colonies.

Discussion
The presented results evinced that WingMarks was able to retrieve the coordinates of wing vein junctions in *Apis mellifera* in a facile manner, and those coordinates could be successfully used to discriminate between different genetic variations of this species. To the best of our knowledge, WingMarks would be the only free software available for the highly precise, repeatable and automatic recognition of the 19 wing landmarks of vein intersections in 100% of the wings using different equipment for image acquisition. The new method does not require previous experience and can be used by non-specialists. The only limitation of the new tool would be the need to use a template for each subspecies, although this only has to be defined the first time.

Several attempts have been made to build automatic honey bee wing vein junction identification systems based on image analysis (Meixner et al., 2013; Tofilski, 2004). As noted above, Tofilski (2004) developed DrawWing open software to automatically determine the coordinates of eighteen wing landmarks. Using the equipment described herein, however, DrawWing failed over 90% of the times.
As compared to DrawWing, WingMarks software may be deemed as a more versatile and accurate alternative to automatically determine landmark coordinates for honeybee forewing venation, and it does not require technical expertise. Furthermore, on the few occasions when it failed, the automatically determined landmark was close to the actual vein intersection, so a small shift was enough to correct it. This is a differential feature as compared to DrawWing, which—in addition to having a high failure rate—returned all the incorrect landmarks in a corner, making the correction more time-consuming than manually plotting the points directly in tpsDig.

Apropos of ApiClass website (Baylac et al., 2008), despite having been used in some wing morphometric studies (Barour et al., 2011), it is worth noting that its current usability is severely limited by the fact that it no longer returns coordinate values and that it does not allow the user to edit the coordinates in case of error. In terms of precision, when high quality images (from Leica microscope) were used, the automatic detection of the 19 wing landmarks was similar for both tools. On the other hand, WingMarks performed better than ApiClass when lower quality images (from Celestron microscope) were used, as it automatically detected 100% of the landmarks in a higher proportion of those images. In addition to its higher tolerance to lower image quality, the possibility of correcting the landmarks in case of error in a fast and easy way, allowing the analysis of all images, may be regarded as a substantial advantage.

The high repeatability of landmark determination using WingMarks software would also be a relevant feature, since it avoids the need to repeat wing measurements in order to reduce measurement error, as proposed by different authors (Miguel et al., 2011; Rasic et al., 2015). Furthermore, its tolerance to rotation and displacement of the wing in the image would increase the reliability of morphometric determinations for non-expert users.
Landmark coordinates provided by the WingMarks software were successfully used in the second experiment to correctly classify three different subspecies and a hybrid of *Apis mellifera* after applying geometric morphometrics with the MorphoJ free-software.

A database of wing diagrams is currently under preparation in order to automatically perform the classification of new subjects, skipping the template preparation step.

Although WingMarks is designed to extract landmarks from honey bees wing automatically, it may be easily adapted to analyze wings from other insect species just by changing the parameter that defines the number of landmarks. It may also be readily modified/improved by other authors, as its source code is available upon request.

It is also worth noticing that the performance of WingMarks was found to be slightly better when the code was run directly under MATLAB environment than for the standalone version. Such difference may be partly ascribed to the alternative algorithm that had to be used in the standalone version to replace the control point selection tool (CPSELECT) function, which cannot be deployed using MATLAB Compiler.

**Conclusions**

The new software presented herein (WingMarks) constitutes a versatile, precise and accurate tool for the automatic recognition of wing vein junctions of *Apis mellifera*, allowing a rapid identification of bee genetic diversity using geometric morphometrics.

Finally, the Creative Commons (CC BY-NC-SA 4.0) license under which the program is distributed allows that this tool can be adapted to different insect species.
**Funding details**

This work was supported by Instituto de Estudios Altoaragoneses under “XXXI Concurso de Ayudas de Investigación 2015” program; and Ibercaja Obra Social under “Proyectos de investigación, desarrollo e innovación para jóvenes investigadores” program.

**Disclosure statement**

The authors declare no conflict of interest.

**References**

Barour, C., Tahar, A., & Baylac, M. (2011). Forewing Shape Variation in Algerian Honey Bee Populations of *Apis mellifera intermissa* (Buttel-Reepen, 1906) (Hymenoptera: Apidae): A Landmark-Based Geometric Morphometrics Analysis. *African Entomology, 19*(1), 11-22. doi: 10.4001/003.019.0101

Baylac, M., Garnery, L., Tharavy, D., Pedraza-Acosta, J., Rortais, A., & Arnold, G. (2008). *Apiclass, an automatic online wing morphometric expert system for honey bee worker identification*. Retrieved from http://apiclass.mnhn.fr/

Charistos, L., Hatjina, F., Bouga, M., Mladenovic, M., & Maistros, A. D. (2014). Morphological Discrimination of Greek Honey Bee Populations Based on Geometric Morphometrics Analysis of Wing Shape. *Journal of Apicultural Science, 58*(1), 75-84. doi: 10.2478/jas-2014-0007

De la Rúa, P., Jaffê, R., Dall'Olio, R., Muñoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honeybees. *Apidologie, 40*(3), 263-284. doi: 10.1051/apido/2009027

Dolati, L., Rafie, J. N., & Khalesro, H. (2013). Landmark-Based Morphometric Study in the Fore and Hind Wings of an Iranian Race of European Honeybee (*Apis*...
Francoy, T. M., Grassi, M. L., Imperatriz-Fonseca, V. L., de Jesús May-Itzá, W., & Quezada-Euán, J. J. G. (2011). Geometric morphometrics of the wing as a tool for assigning genetic lineages and geographic origin to Melipona beecheii (Hymenoptera: Meliponini). *Apidologie, 42*(4), 499-507. doi: 10.1007/s13592-011-0013-0

Francoy, T. M., Wittmann, D., Drauschke, M., Muller, S., Steinhage, V., Bezerra-Laure, M. A. F., . . . Goncalves, L. S. (2008). Identification of Africanized honey bees through wing morphometrics: two fast and efficient procedures. *Apidologie, 39*(5), 488-494. doi: DOI 10.1051/apido:2008028

Francoy, T. M., Wittmann, D., Steinhage, V., Drauschke, M., Muller, S., Cunha, D. R., . . . Goncalves, L. S. (2009). Morphometric and genetic changes in a population of *Apis mellifera* after 34 years of Africanization. *Genetics and Molecular Research, 8*(2), 709-717

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources, 11*(2), 353-357. doi: 10.1111/j.1755-0998.2010.02924.x

Meixner, M. D., Pinto, M. A., Bouga, M., Kryger, P., Ivanova, E., & Fuchs, S. (2013). Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. *Journal of Apicultural Research, 52*(4), 1-28. doi: 10.3896/Ibra.1.52.4.05

Miguel, I., Baylac, M., Iriondo, M., Manzano, C., Garnery, L., & Estonba, A. (2011). Both geometric morphometric and microsatellite data consistently support the differentiation of the *Apis mellifera* M evolutionary branch. *Apidologie, 42*(2), 150-161. doi: 10.1051/apido/2010048
Nawrocka, A., Kandemir, İ., Fuchs, S., & Tofilski, A. (2017). Computer software for identification of honey bee subspecies and evolutionary lineages. *Apidologie.* doi: 10.1007/s13592-017-0538-y

Oleksa, A., & Tofilski, A. (2015). Wing geometric morphometrics and microsatellite analysis provide similar discrimination of honey bee subspecies. *Apidologie,* 46(1), 49-60. doi: 10.1007/s13592-014-0300-7

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica,* 11(285-296), 23-27

Rasic, S., Mladenovic, M., & Stanisavljevic, L. (2015). Use of geometric morphometrics to differentiate selected lines of Carniolan honeybees (Apis mellifera carnica) in Serbia and Montenegro. *Archives of Biological Sciences,* 67(3), 929-934. doi: 10.2298/abs140224054r

Rohlf, F. J. (2001). TpsDig: digitize landmarks from image files, scanner, or video. Stony Brook, NY, USA: Department of Evolutionary Biology, University of New York.

Smith, D. R., Crespi, B. J., & Bookstein, F. L. (1997). Fluctuating asymmetry in the honey bee, Apis mellifera: effects of ploidy and hybridization. *Journal of Evolutionary Biology,* 10(4), 551-574

Steinhage, V., Arbuckle, T., Schröder, S., Cremers, A. B., & D., W. (2001). *ABIS: Automated Identification of Bee Species.* Retrieved from BIOLOG Workshop, German Programme on Biodiversity and Global Change, Status Report.

Tofilski, A. (2004). DrawWing, a program for numerical description of insect wings. *J Insect Sci,* 4, 17
Figure 1. (a) numbering of the vein junctions; (b) selection of two points for tilt correction; (c) greyscale image; (d) output screen for user verification and/or correction of the automatically-detected landmarks.

Figure 2. Scatter plot of three honeybee subspecies (A. m. iberiensis, blue; A. m. ligustica, violet; and A. m. carnica, green) and of Buckfast hybrid (red) based on CVA for Cartesian coordinates of the landmarks on the right forewing discrimination of colonies.