A statistical approach for spectral measure of color variation of black - orange - black (BOB) pattern in small parasitic wasps

Marcela Alfaro-Córdoba, Rebeca Mora-Castro, Marcela Hernández-Jiménez, Esteban Avendaño-Soto, Paul Hanson-Snortum

Universidad de Costa Rica

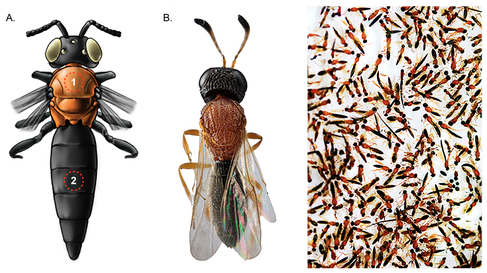

Figure 1: Black - orange - black (BOB) pattern in small parasitic wasps (Hymenoptera: Scelionidae). Photo credit: Rebeca Mora

Are the colors different?

Differences among orange and black tones, can lead us to biological hypothesis about prey behavior, differences between species among genera and also, pose questions about taxonomy.

Why is so hard to distinguish between colors?

Human eye does not necessarily perceives the color in the same way as other animals. Hence, measuring the color using microspectrophotometry can gives us more clues about color differences.

Why is this interesting in data science and statistics?

- Reproducible data workflow from the measurements to the visualization and analysis.
- Testing color differences both inter and intragenera using Functional Data Analysis.
- Comparing results with other methods used before: statistical analysis of mean reflectance and Euclidean distances for color components (L, a, b).
- As a confirmatory analysis, a study of color differences using reflectance spectra separated into spectral color component contributions (Red, Green and Blue) is performed.

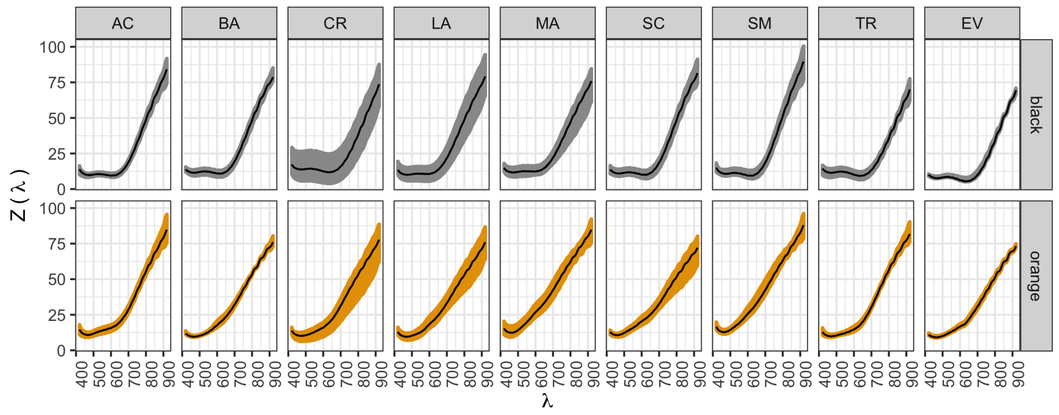

Figure 2: 1185 recorded points: the reflectance (Z) as a function of wavelength value (λ), where λ is a sequence of 1185 equidistant points from 420.2 nm to 919.7 nm. 4 specimens for each of the 8 Scelionoidae genera, and 4 specimens for *Evaniella*.

Methods

Statistical test for differences of colors depending on the measure:

- The **univariate measure**: m_{ijk} difference in a pair of for k represents repetitions, j specimens, and i genera.
- The **functional measure**: $Z_{ijk}(\lambda)$ was measured for each wavelength λ, and repeated k times in each specimen j, and genera i.
- The **multivariate distance measure**: a Euclidean distance $\Delta E = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2}$ for each pair of curves being compared.

Statistical Result

Functional Data Analysis (FANOVA) of color comparisons gives the most robust measure of difference between reflectance curves. Results are consistent with the multivariate method, but not with the univariate, which fails to find differences even in orange vs black colors.

RGB Color Components

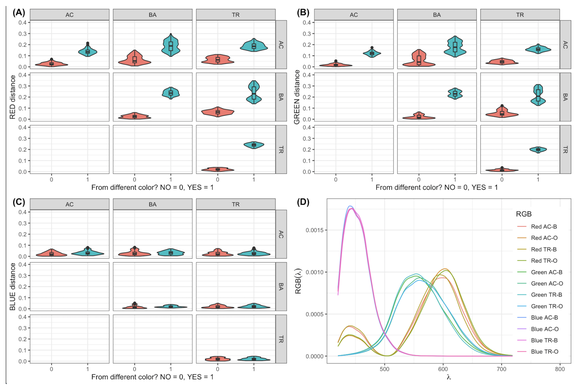

Figure 3: Distributions of curve distance for (A) green, (B) red and (C) blue components, calculated for pairs of curves of three different genera: *Acanthoscelio* (AC), *Baryconus* (BA) and *Triteleia* (TR), and two scenarios: when they have the same color (in vermilion) and when their colors differ (bluish green). Plot (D) is an example of how the R, G, and B curves look for *Acanthoscelio* (AC) and *Triteleia* (TR) in orange and black spots.

Conclusion

While univariate comparisons can be useful for general description, the most recommended method to test for color differences with microspectrophotometer data is using FANOVA, and controlling factors like: repetitions, spots, specimens and genera. Multivariate distance measures can be helpful to characterize differences among color components (blue, red, green), using violin plots to describe those differences. The RGB spectral components method is proposed for studying the origin of the color in the cases when a direct chemical composition is not possible. In the case of the genera studied, the spectral blue components of the orange and black color were found to be almost identical, suggesting that there is a common compound for the pigments.

References

[1] David Causeur, Ching-Fan Sheu, Mei-Chen Chu, and Flavia Rufini.
ERP: Significance analysis of event-related potentials data, 2018.
R package version 2.1.

[2] Tomasz Gorecki and Lukasz Smaga.
fdANOVA: Analysis of variance for univariate and multivariate functional data, 2018.
R package version 0.1.2.

[3] Hadley Wickham.
tidyverse: Easily Install and Load the 'Tidyverse', 2017.
R package version 1.2.1.

Contact Information

- Web: https://malfaro.netlify.com/
- Email: marcela.alfarocordoba@ucr.ac.cr

Affiliations

CICIMA, CIMPA, Escuela de Biología, Escuela de Física, Universidad de Costa Rica.

R packages: R packages erpFtest [1], fdANOVA [2] for FDA. Data formatting and figures Tidyverse [3].
Pre-print: https://www.biorxiv.org/content/10.1101/652594v1
Repo with code and data: https://github.com/malfaro2/mora$_et_al$

UNIVERSIDAD DE COSTA RICA