**Introduction:**

In this investigation, we used data from an observational study centered on the morphological dimensions of Tephritis conura across distinct geographic regions situated on the eastern and western peripheries of the Baltic Sea. This analysis primarily examines the influence of regional categorization (west/east) and host plant distinctions (Oleraceum, Heterophyllum) on allopatric and sympatric conditions, particularly emphasizing their possible interactive impact on the body length of Tephritis conura individuals.

**Methods:**

In the initial phase of our investigation, we undertook an exploratory analysis of the response variable, body length, to gain insights into the underlying data distribution. Therefore, we constructed a histogram to visualize the distribution pattern of body length within the Tephritis population.

Subsequently, we explored the variation in body length distribution across different regions, host plants, and allopatric and sympatric individuals. This step provided additional information to our understanding of data distribution, revealing subtle trends within the dataset.

We applied various linear and mixed-effect models to discern the most suitable model for our research question. A total of six linear models and one mixed-effect model were assessed and evaluated. The model selection was guided by the Akaike Information Criterion (AIC).

Having identified the best-fitting model, we analyzed the variance (ANOVA) to quantify the statistical differences among Tephritis groups based on their region, host plant, and allopatric or sympatric individual classifications. For analysis of the main effects and interactions, we employed the effect size eta-squared (η²), calculated as the ratio of the sum of squares for a given effect to the total sum of squares.

Finally, after the analysis of the model, we used Interaction plots to communicate the outcomes of our ANOVA analysis effectively. These visualizations act as effective tools to demonstrate the main effects of region and host plants on the body length of Tephritis.

**Results:**

Our analysis initiation involved assessing the distribution of the response variable (body length). The histogram revealed a tendency towards a normal distribution, facilitating the application of multiple subsequent analytical tools.

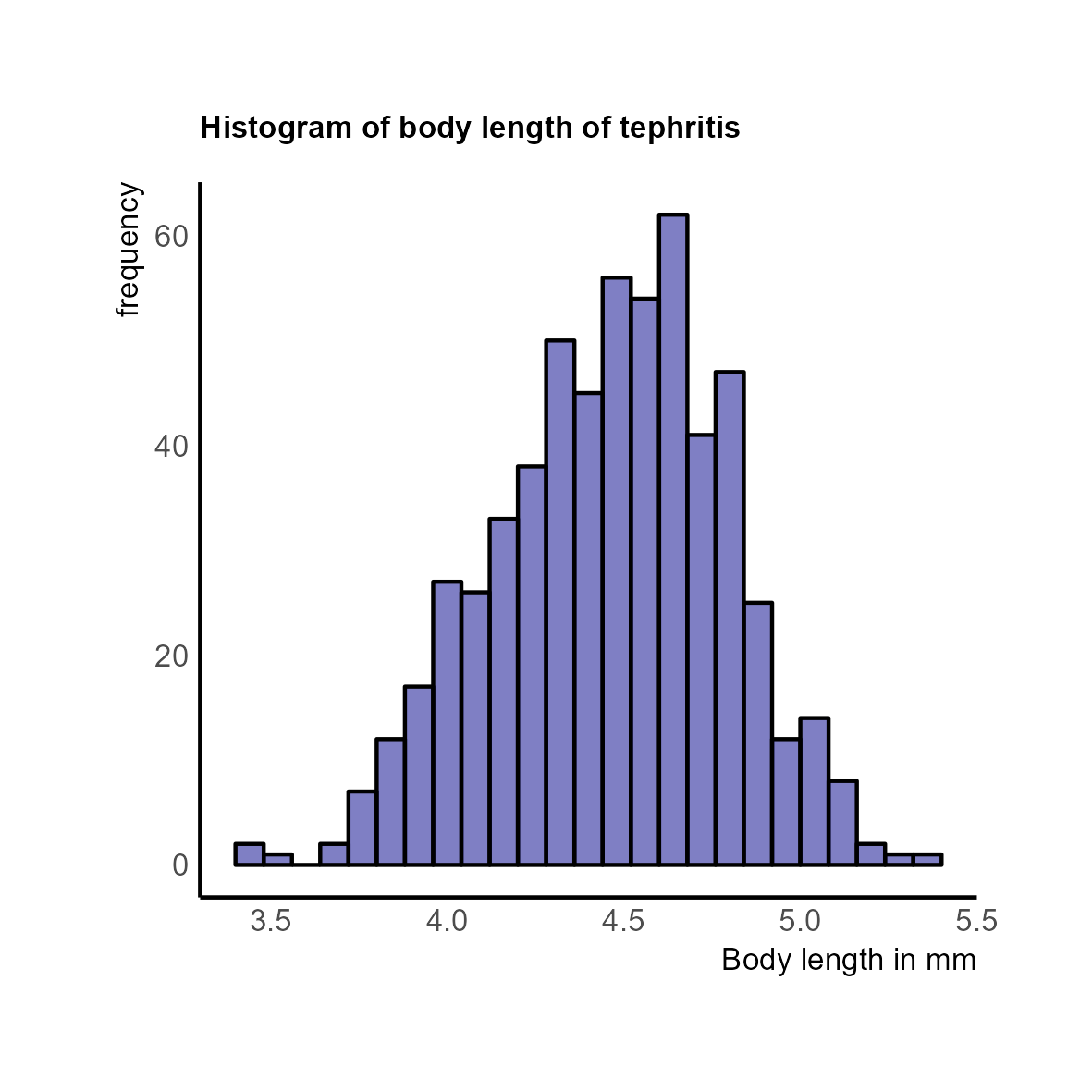


Figure 1: Histogram of body length of tephritis in the dataset.

We examined the variability in the distribution of body length across diverse regions, host plants, and individuals in allopatric and sympatric settings. Initial observations revealed a smaller size of Tephritis hosted on Oleraceum in the western region than other Tephritis specimens (Appendix Figure 4A). Subsequent analysis indicated that the interaction between host plants and patry did not exhibit a discernible visual trend in the density plot (Appendix Figure 4B). Notably, the data unveiled that allopatric individuals appeared smaller in both the West and East regions (Appendix Figure 4C). To conclude, a comprehensive exploration of boxplots encompassing all conceivable combinations of regions, host plants, and allopatric and sympatric individuals highlighted a consistent trend: individuals bred on Oleraceum in the West region appeared smaller than their counterparts in the East (Figure 2). An Overview on the eight possible group is given in table 1.

Table : Table of possible groups of combination of Baltic, Hostplant and Patry

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Group name | | Baltic | Hostplant | Patry |
| EOS | East | | Oleraceum | Sympatric |
| EOA | East | | Oleraceum | Allopatric |
| EHS | East | | Heterophyllum | Sympatric |
| EHA | East | | Heterophyllum | Allopatric |
| WOS | West | | Oleraceum | Sympatric |
| WOA | West | | Oleraceum | Allopatric |
| WHS | West | | Heterophyllum | Sympatric |
| WHA | West | | Heterophyllum | Allopatric |

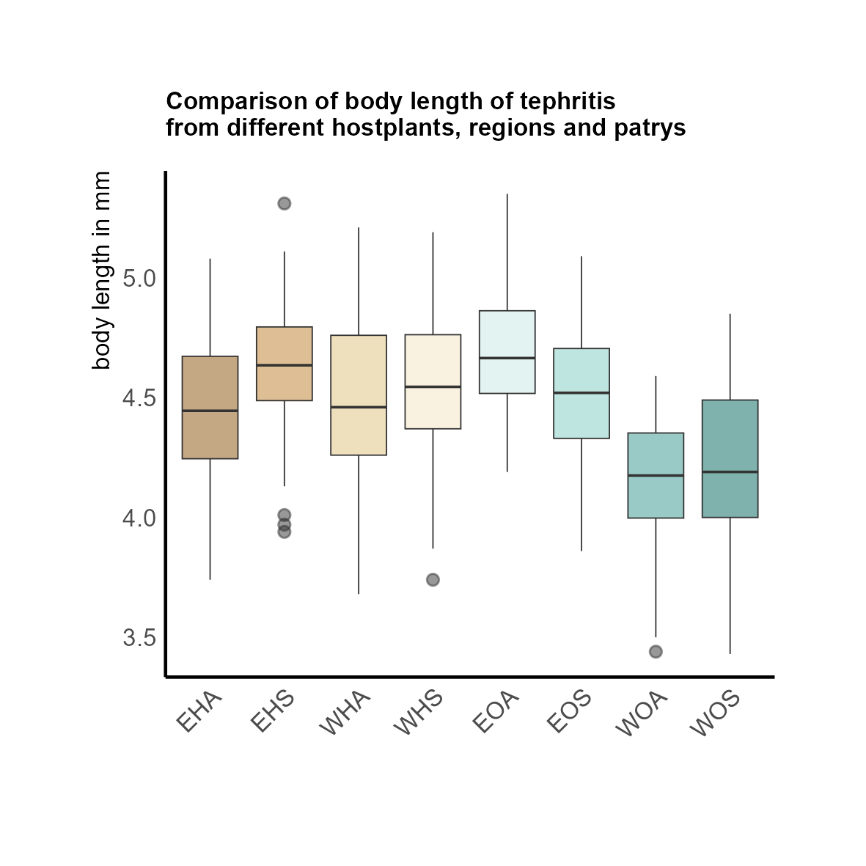


Figure 2: Distribution of body length in each of the eight possible groups.

In a systematic examination, diverse linear and mixed-effect models were thoroughly assessed to identify the optimal model aligning with the objectives of our research inquiry. An evaluation, encompassing six linear models and one mixed-effect model, led to the determination that a linear model incorporating all variables (Hostplant, Baltic, Patry) and their interactions, demonstrated superior compatibility with our dataset.

Table : Comparison of models with different modification of variables

|  |  |  |  |
| --- | --- | --- | --- |
| Model | AIC | logLik | weight |
| lm(BL ~ Hostplant\*Baltic\*Patry) | 209.5692 | -103.78459 | 1 |
| lm(BL ~ Hostplant\*Baltic+Patry) | 222.4416 | -119.22080 | 0 |
| lmer(BL ~ 1 +(1|HostplantBalticPatry)) | 275.4534 | -132.72672 | 0 |
| lm(BL ~ Hostplant\*Patry+Baltic) | 318.2449 | -132.72672 | 0 |
| lm(BL ~ Patry\*Baltic+Hostplant) | 319.8969 | -153.94846 | 0 |
| lm(BL ~ Baltic) | 342.9737 | -168.48684 | 0 |
| lm(BL ~ 1) | 347.2582 | -171.62912 | 0 |

Hence, our analysis centered on categorical variables, with the chosen model serving as the framework for conducting an analysis of variance (ANOVA). This statistical approach allowed for the distinctions of significant differences among diverse groups. To further increase the comprehensiveness of our analysis, the calculated effect size was incorporated into the ANOVA table. Specifically, for main effects and interactions, we employed eta-squared (η²), calculated as the ratio of the sum of squares for a given effect to the total sum of squares. This measure offers an estimation of the proportion of variance in the dependent variable attributed to a specific effect, enhancing our understanding of the impact of our variables and their interactive effects on the observed outcomes.

The findings reveal meaningful associations between certain factors and body length. Variations in Hostplant are linked to differences in body length, suggesting that the type of Hostplant plays a role in determining body length. Similarly, differences in Baltic regions are strongly associated with variations in body length.

The interaction between Hostplant and Baltic, as well as Hostplant and Patry, showed significant effects, indicating that the combined influence of these variables surpasses the sum of their individual impacts on body length. Also, the three-way interaction between the variables showed significant effect. This suggests a complex interplay between Hostplant, Baltic, and Patry, with a particularly notable three-way interaction effect.

However, the residuals, representing unexplained variance in body length after considering these factors, indicate that other aspects mainly contribute to body length variations. The substantial residual variations suggest a significant portion of body length remains unpredictable by Hostplant, Baltic, and Patry, emphasizing a high general variability in body length.

Notably, Patry and the interaction between Patry and Baltic did not show significant effects on body length. The evidence does not support the conclusion that differences in Patry or a joint effect of Baltic and Patry significantly contribute to variations in body length.

Table 3: ANOVA of the selected model with body length as response variable

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Analysis of Variance Table (responds variable = body length) | | | | | |
|  | Df | Sum Sq | η² | F value | Pr(>F) |
| Hostplant | 1 | 0.659 | 0.01 | 7.7766 | 0.0055 |
| Baltic | 1 | 4.143 | 0.08 | 48.8569 | 7.65e-12 |
| Patry | 1 | 0.023 | 4.77e-4 | 0.2745 | 0.6005 |
| Hostplant:Baltic | 1 | 5.274 | 0.10 | 62.1908 | 1.57e-14 |
| Hostplant:Patry | 1 | 1.384 | 0.03 | 16.3163 | 6.09e-05 |
| Baltic:Patry | 1 | 0.144 | 2.94e-3 | 1.6967 | 0.1932 |
| Hostplant:Baltic:Patry | 1 | 1.119 | 0.02 | 13.1984 | 0.0003 |
| Residuals | 575 | 48.758 | 0.76 |  |  |

The influence of Baltic regions and Hostplants on body length differences becomes evident when observing interaction plots. In general, individuals tend to be smaller in the west region compared to the east, with a notable effect for those bred on Oleraceum. However, allopatric individuals bred on Heterophyllum in the west region appear larger (Figure 3A). When focusing solely on the hostplant effect, individuals bred on Oleraceum are generally smaller than those bred on Heterophyllum. Interestingly, allopatric individuals in the east region show a significant size difference, being larger when bred on Oleraceum rather than Heterophyllum (Figure 3B).

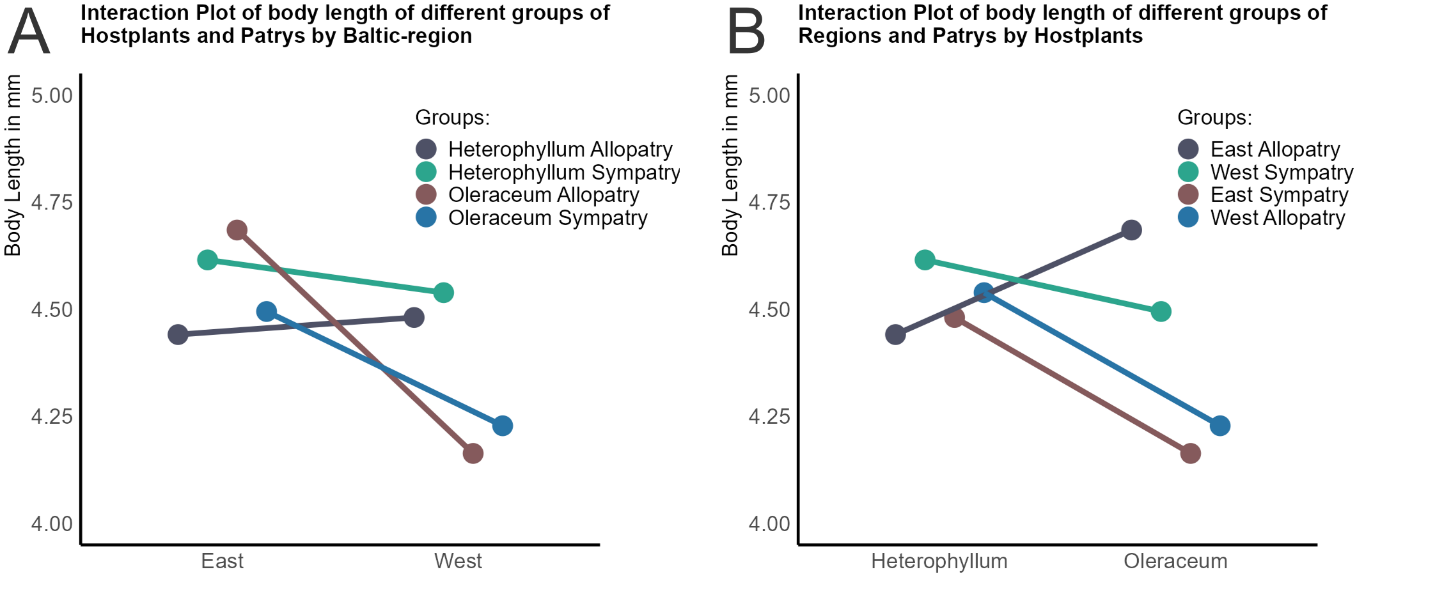


Figure : Interaction Plot. (A) Effect of Baltic on the difference in body length. (B) Effect of Hostplants on the body length

**Conclusion:**

In conclusion, our investigation into the body size of Tephritis conura across diverse Baltic regions, host plants, and patry conditions yielded significant findings. The observed variations in body length associated with different host plants and Baltic regions underscore their role in determining the body size of Tephritis conura. This connection may be attributed to the varying nutritional resources provided by distinct host plants and the environmental conditions prevalent in different Baltic regions.

The interactive effects observed between host plants and Baltic regions, as well as host plants and patry conditions, add complexity to these relationships, suggesting that the combined influence of these variables surpasses the sum of their individual impacts. This interplay may be indicating adaptive responses to environmental influences, where certain combinations of host plants and Baltic regions create more favorable conditions for the development of Tephritis than others.

The lack of significant impact from patry conditions and the interaction between patry and Baltic regions on body length suggests that these factors may not play a substantial role in shaping the observed variations.

Observations from interaction plots further emphasize the pronounced influence of Baltic regions and host plants on body length differences. The smaller sizes observed in individuals bred on Oleraceum in the west region may be linked to the availability of specific nutrients or environmental conditions in that habitat. Notable exceptions for allopatric individuals bred on Heterophyllum, appearing larger in the west region, might be indicative of favorable conditions provided by this particular host plant in that region.

In contrast, a consistent pattern of smaller body sizes in individuals bred on Oleraceum, as opposed to those bred on Heterophyllum, is observed, with an exception noted for allopatric individuals in the east region where the trend is reversed. This difference in body size may be attributed to variations in the nutrient composition of Oleraceum and Heterophyllum, influencing the availability of resources Tephritis conura. It is possible that allopatric individuals in the east region have developed a better adaptation to the specific nutrient composition of Oleraceum, potentially contributing to the observed increase in body size within this particular group.

In essence, our study highlights the nature of the factors influencing the body size of Tephritis conura, offering valuable insights into the complex interactions between Baltic regions, host plants, and patry conditions.

**Appendix:**

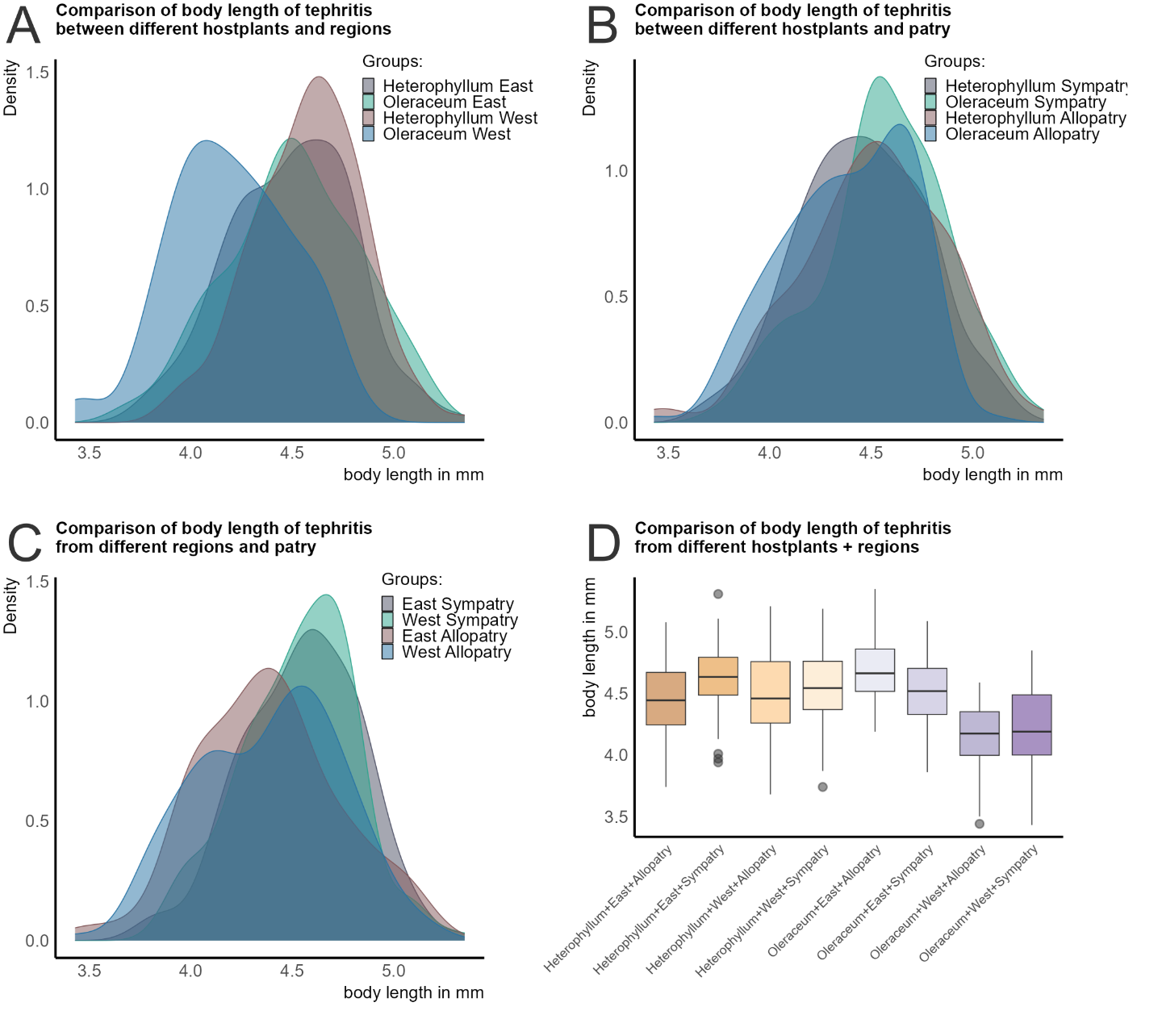


Figure 4: Density plots illustrating the distribution of body length in Tephritis between different hostplants and region (A), hostplants and patry (B), and regions and patry (C). Figure D shows the distribution of body length in each of the eight possible groups.

**Link to the R-Code:**

https://github.com/JaSt17/BIOS4/blob/main/exercises/MidTerm.R