ANOVA of Host Effects in Pieris Butterflies

The Pieris genus of Lepidoptera butterflies preferentially oviposit and feed on plant species of the Brassicaceae family, notably Barbarea and Berteroa species. This study collected data from 287 individual F1 butterflies (observing the mother generation as well as the offspring), noting the host plant that the mother emerged from as well as the host plant the offspring was laid upon.

The aim of this analysis is to determine and quantify host effects of both the larvae and mother on 1) development time, 2) growth rate, and 3) adult weight. We hypothesize that the host of the larval generation is more significant than the maternal host in all offspring effects. We also hypothesize that adult weight will not be significantly different between hosts.

**Methods/Analysis**

With the collected data from 287 F1 individuals (n=287), we first sorted samples into 4 groups based on larval and maternal hosts to determine balance in the sampling and check for potential bias among mothers to oviposit on their natal host plant. Of the 287 samples, 102 (35.54%) individuals laid on Barbarea came from mothers also laid on Barbarea; 78 (27.18%) laid on Barbarea came from mothers laid on Berteroa; 55 (19.16%) laid on Berteroa came from mothers laid on Barbarea; and 52 (18.12%) laid on Berteroa came from mothers also laid on Berteroa.

For this analysis, since the host pairings in the sampling were unbalanced and multiple independent factors are being analyzed, data analysis was performed using an unbalanced two-way Anova model in R Studio software. We first fit a linear model for each of the effects (development time, growth rate, and adult weight) against the larval-maternal host pairings.

> summary(lm(DevelopmentTime ~ LarvalHost \* MaternalHost, data=dat))

> summary(lm(GrowthRate ~ LarvalHost \* MaternalHost, data=dat))

> summary(lm(AdultWeight ~ LarvalHost \* MaternalHost, data=dat))

Then the Anova two-way unbalanced model function in R was used to perform the regression analysis for the significance of the variation within and between host groups on the measured effects. This required using the “car” package in R.

> library(car)

> mod2 <- lm(DevelopmentTime ~ LarvalHost \* MaternalHost, data=dat)

> Anova(mod2)

> mod2 <- lm(GrowthRate ~ LarvalHost \* MaternalHost, data=dat)

> Anova(mod2)

> mod2 <- lm(AdultWeight ~ LarvalHost \* MaternalHost, data=dat)

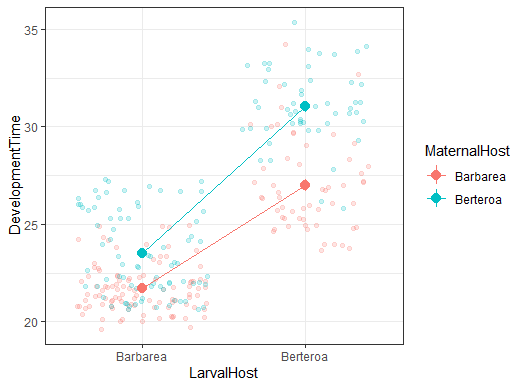
> Anova(mod2)

**Results**

Sampling, while more biased towards Barbarea host individuals, seemed to indicate low preference of mothers for oviposition on the same host (Fig. 1), as nearly half of the mothers sampled from each host group laid their larvae on the opposite host group (78/180=43.33% Barbarea laid by Berteroa mothers and 55/107=51.40% Berteroa laid by Barbarea mothers). This may indicate that Pieris butterflies on Barbarea are simply more abundant in the sampling area but no significant preference exists for one host over the other.

The mean development time for Pieris individuals was lowest in the Barbarea-Barbarea group (m=21.69608) (Fig. 2) and the between group variation showed greatest significance for this effect, as compared to growth rate and adult weight; the F value of LarvalHost:MaternalHost in the development time Anova (F=23.05, Table 2) is higher than both other effects.

As expected, the larval host had a greater effect than maternal host on all observed effects in offspring. Larvae laid on Barbarea showed lower development time, faster growth rate, and marginally higher adult weight than those laid on Berteroa. While development time showed greatest significance between larval and maternal hosts, growth rate showed steeper variation between larval hosts but less variation between maternal hosts.

A graph with red and blue dots

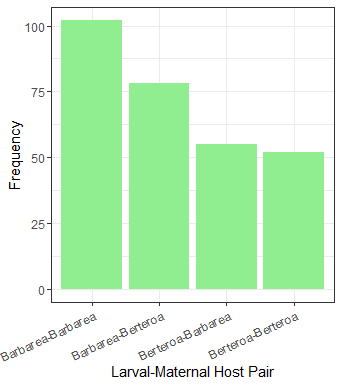
Description automatically generated

Figure 3: Development Regression

Figure 2: Hosts vs Development Time

Figure : Sampling distribution

A diagram of a graph

Description automatically generated with medium confidenceA graph with red and blue dots

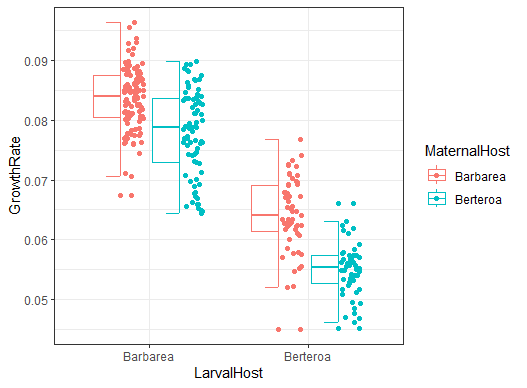
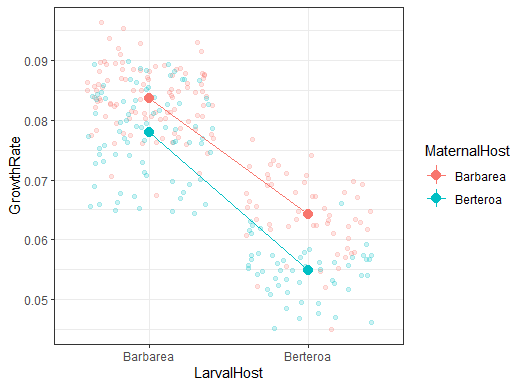
Description automatically generated

Figure 7: Adult Weight Regression

Figure 6: Hosts vs Adult Weight

Figure 5: Growth Rate Regression

Figure 4: Hosts vs Growth Rate

**Table 1: Linear Regression Error and Significance of Development Time**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Coefficients: |  |  |  |  |  |
|  | Estimate | Std. Error | t value | Pr(>|t|) |  |
| (Intercept) | 21.6961 | 0.1854 | 117.033 | < 2e-16 | \*\*\* |
| LarvalHostBerteroa | 5.3039 | 0.3132 | 16.934 | < 2e-16 | \*\*\* |
| MaternalHostBerteroa | 1.8167 | 0.2816 | 6.451 | 4.80e-10 | \*\*\* |
| LarvalHostBerteroa:MaternalHostBerteroa | 2.2025 | 0.4588 | 4.801 | 2.56e-06 | \*\*\* |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |

Residual standard error: 1.872 on 283 degrees of freedom

Multiple R-squared: 0.7734, Adjusted R-squared: 0.771

F-statistic: 322.1 on 3 and 283 DF, p-value: < 2.2e-16

**Table 2: Unbalanced model two-way ANOVA of Development Time**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Df | F value | Pr(>F) |  |
| LarvalHost | 2682.41 | 1 | 765.21 | < 2.2e-16 | \*\*\* |
| MaternalHost | 496.87 | 1 | 141.74 | < 2.2e-16 | \*\*\* |
| LarvalHost:MaternalHost | 80.80 | 1 | 23.05 | 2.561e-06 | \*\*\* |
| Residuals | 992.05 | 283 |  |  |  |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | |  |

**Table 3: Linear Regression Error and Significance of Growth Rate**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Coefficients: | | | | |  |
|  | Estimate | Std. Error | t value | Pr(>|t|) |  |
| (Intercept) | 0.0837158 | 0.0005897 | 141.952 | < 2e-16 | \*\*\* |
| LarvalHostBerteroa | -0.0194909 | 0.0009964 | -19.561 | < 2e-16 | \*\*\* |
| MaternalHostBerteroa | -0.0057776 | 0.0008959 | -6.449 | 4.85e-10 | \*\*\* |
| LarvalHostBerteroa:MaternalHostBerteroa | -0.0035335 | 0.0014594 | -2.421 | 0.0161 | \* |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |

Residual standard error: 0.005956 on 283 degrees of freedom

Multiple R-squared: 0.7763, Adjusted R-squared: 0.774

F-statistic: 327.4 on 3 and 283 DF, p-value: < 2.2e-16

**Table 4: Unbalanced model two-way ANOVA of Growth Rate**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Df | F value | Pr(>F) |  |
| LarvalHost | 0.0299064 | 1 | 843.0077 | <2e-16 | \*\*\* |
| MaternalHost | 0.0035847 | 1 | 101.0475 | <2e-16 | \*\*\* |
| LarvalHost:MaternalHost | 0.0002080 | 1 | 5.8621 | 0.0161 | \* |
| Residuals | 0.0100397 | 283 |  |  |  |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |

**Table 5: Linear Regression Error and Significance of Adult Weight**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Coefficients: | | | | | |
|  | Estimate | Std. Error | t value | Pr(>|t|) |  |
| (Intercept) | 65.3549 | 0.9386 | 69.629 | < 2e-16 | \*\*\* |
| LarvalHostBerteroa | -11.8513 | 1.5858 | -7.473 | 9.81e-13 | \*\*\* |
| MaternalHostBerteroa | 1.3695 | 1.4259 | 0.960 | 0.3377 |  |
| LarvalHostBerteroa:MaternalHostBerteroa | -4.4962 | 2.3227 | -1.936 | 0.0539 | . |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |

Residual standard error: 9.48 on 283 degrees of freedom

Multiple R-squared: 0.3456, Adjusted R-squared: 0.3387

F-statistic: 49.82 on 3 and 283 DF, p-value: < 2.2e-16

**Table 6: Unbalanced model two-way ANOVA of** **Adult Weight**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Df | F value | Pr(>F) |  |
| LarvalHost | 13019.8 | 1 | 144.8883 | <2e-16 | \*\*\* |
| MaternalHost | 7.5 | 1 | 0.0833 | 0.7731 |  |
| LarvalHost:MaternalHost | 336.7 | 1 | 3.7471 | 0.0539 | . |
| Residuals | 25430.7 | 283 |  |  |  |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | | | | |