

Effect of different population of flower on the gland area, with the gland-stigma distance as the associated factor

Pham Xuan Huy Nguyen

December 2022

1 Background

Reproduction is one of the most important part in a life cycle of a plant, and the phenotypes of flower in floral plants to attract pollinator of plants are what evolutionary botanists are interested in. The combination of different patterns among floral parts can affect the pollination process of pollinators. 9 populations of flowers in Costa Rica are selected and 9 measurements of floral part are made. In this research, we will investigate the effect of different population of flower on the gland area (GA), with the gland-stigma distance (GSD) as the associated factor. The GSD tells which size of pollinator will fit with the flower, while GA refers to the reward for the pollinator.

2 Methods and Results

Homogeneity of regression slopes

The following lines read the file data and plot the homogeneity of regression slopes. In Figure 1, each colours represents the data from a particular measurement in GA and GSD, and the colours tell us the population. The lines are the regression slopes for the particular population. There are unequal positive relationships between GA and GSD in all populations except S20, which appears to be no relationship at all, giving doubt whether there is homogeneity of regression slopes. Many points fall outside the grey area, which indicate that the homogeneity may be violated.

```
blossoms = read.csv('blossoms.csv')
library(ggplot2)
ggplot(blossoms, aes(GSD, GA, colour=pop)) +
  geom_point() + stat_smooth(method=lm) +
  labs(title='Scatter plot of GA and GSD from different populations',
x="Gland-stigma Distance (mm)", y="Gland Area (mm^2)")
```

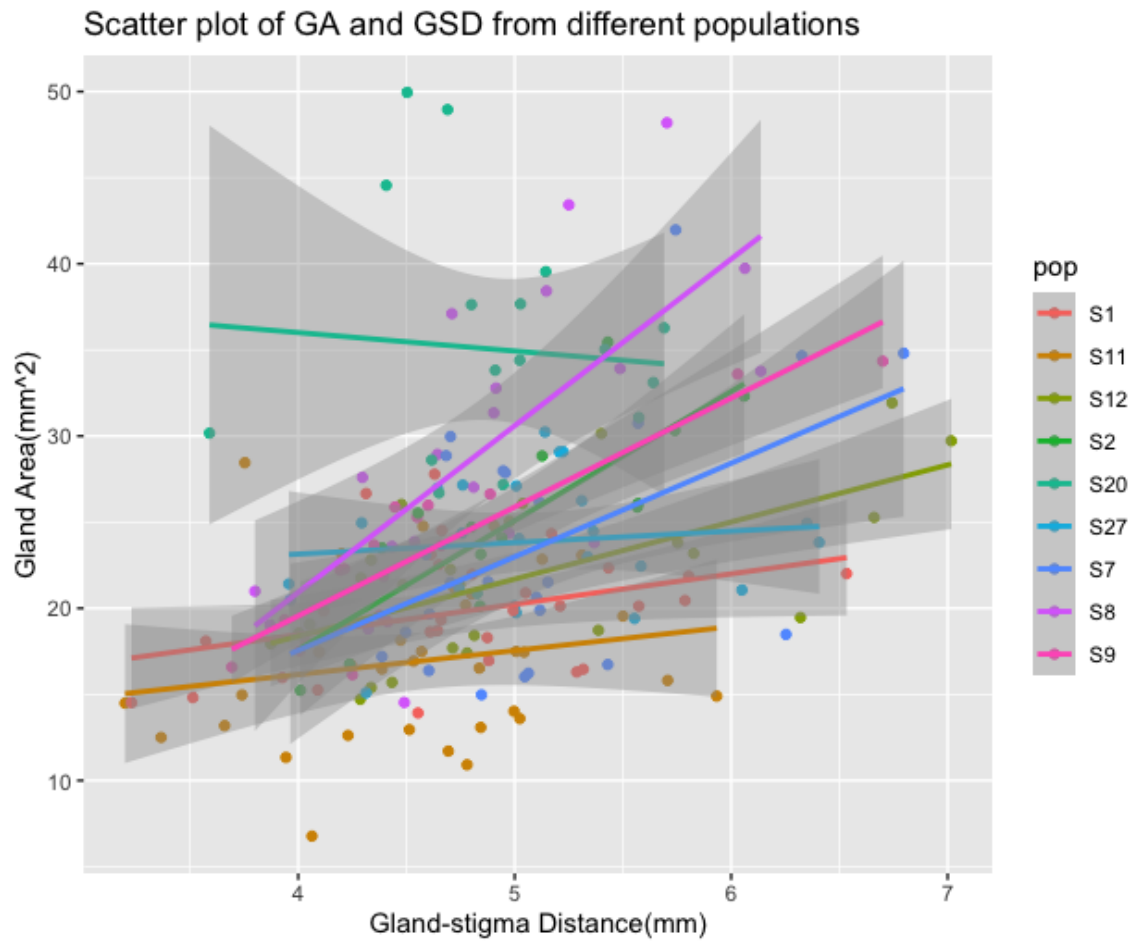


Figure 1: Display the relationship between GA and GSD for 9 populations (different colours).

ANCOVA

The following lines show the ANCOVA results of GA in different population with GSD as predictor and the interaction of population and GSD.

```
library(car)
md = aov(GA ~ GSD + pop + pop:GSD, data=blossoms)
Anova(md, type='III')
```

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	86.0	1	3.6009	0.0592976 .
GSD	47.6	1	1.9922	0.1597820
pop	641.7	8	3.3594	0.0012779 **
GSD:pop	740.2	8	3.8753	0.0003011 ***
Residuals	4440.8	186		

Signif. codes:

0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

The interaction between GSD and population is statistical significant, thus the homogeneity of regression slope is violated. We will proceed assuming the homogeneity of regression slopes assumption is violated. Next, we would like to check for independence of the covariant GSD and the treatment group population.

```
newmd = aov(GSD~pop, data=blossoms)
```

```
summary(newmd)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
pop	8	8.46	1.0577	2.261	0.0248 *
Residuals	195	91.24	0.4679		

Signif. codes:

0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

The model has statistical significant, we will proceed with assumption that we violate the independence assumption between the covariant and the treatment group.

Examine GA over different population when controlling for the GSD

```
mdl = lm(GA ~ GSD + pop, data=blossoms)
anova(mdl)
```

Analysis of Variance Table

Response: GA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
GSD	1	1780.2	1780.18	66.658	4.029e-14 ***
pop	8	4403.4	550.42	20.610	< 2.2e-16 ***
Residuals	194	5181.0	26.71		

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

summary(mdl)

Residuals:

Min	1Q	Median	3Q	Max
-13.8666	-3.4135	-0.1919	2.9980	16.3949

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.5236	2.7382	0.922	0.35787
GSD	3.6349	0.5410	6.719	1.99e-10 ***
popS11	-2.1992	1.2980	-1.694	0.09182 .
popS12	0.9903	1.4069	0.704	0.48233
popS2	4.4178	1.6784	2.632	0.00917 **
popS20	14.6644	1.5711	9.334	< 2e-16 ***
popS27	2.7070	1.4655	1.847	0.06625 .
popS7	2.4400	1.4109	1.729	0.08533 .
popS8	9.2528	1.4938	6.194	3.43e-09 ***
popS9	4.0267	1.5718	2.562	0.01117 *

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 5.168 on 194 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.5441, Adjusted R-squared: 0.523

F-statistic: 25.73 on 9 and 194 DF, p-value: < 2.2e-16

Population S1 has smaller gland area than population S11 on average, and bigger gland area than any other population. From the R-square value, we can conclude this model explain 54.4% of the variance in gland area.

3 Conclusions

There is not enough evidence to show that there is an interaction of population on the gland area with the gland-stigma distance as the associated factor.