

Heritability simulations:

This repository contains R/julia scripts that performs simulations of genetic and environmental effects on phenotypic traits.

The main features of the script include:

Simulation of genetic variation: The script simulates parental genotypes for a specified number of loci, considering both additive and dominant effects. It then generates offspring genotypes based on parental genotypes and calculates resulting phenotypic values. Environmental variation: Environmental effects on phenotypic traits are simulated by specifying a range of environmental values. Phenotypic values are calculated by incorporating both genetic and environmental effects. Regression analysis: Linear regression models are fitted to the simulated data to explore the relationship between parental and offspring phenotypic values. The script calculates regression coefficients, intercepts, and R-squared values to quantify the strength of the relationship. Visualization: The script generates visualizations, including scatter plots of parental-offspring phenotypic values, histograms of regression coefficients and intercepts, and a combined plot displaying regression lines for multiple simulations. The script provides insights into the interaction between genetic and environmental factors in determining phenotypic variation, making it useful for educational purposes and research in quantitative genetics.

How to use it?

1. Download this GitHub repository.
2. Install julia and Pluto.jl. The script also requires an installation of R with the package ggplot2 and gridExtra for visualizing the results.
3. Select the parameters using the sliders below
4. Visualize the scatterplots, the heritability is the slope of the graph

```
RObject{StrSxp}
[1] "C:/Users/alexg/OneDrive/Documents/GitHub-repos/heritability_simulations"

RCall.jl: Learn more about the underlying theory at https://ggplot2-book.org/
```

Each simulation is an activity for the class Ecological Genetics in Stony Brook University.



1 @bind num_simulations Slider(1:500,500,true)

In our classroom, comprising 30 groups of 4 students each, every group is tasked with handling 4 pumpkins. These pumpkins possess genotypes characterized by 6 loci, each consisting of zeroes and ones, dictating the diameter of the pumpkins.

During the initial exercise, students study the concept of additivity within a selection experiment. They start by computing the phenotype of each pumpkin, which involves aggregating the effects of alleles and incorporating environmental variation introduced through a die roll (-3, -2, -1, +1, +2, +3) cm.

Subsequently, students select the pumpkins exhibiting the highest and lowest phenotypes and generate gametes. They accomplish this by assigning values for homozygotes or flipping a coin for heterozygotes at each locus.

These gametes are then combined to produce new offspring, with students calculating the phenotype of each offspring in the same manner.

Finally, the class collaboratively conducts a regression analysis on the midparent and midoffspring averages to extract insights from the experiment.

In the second activity, students replicate the process but adhere to complete dominance rules, where $(1,0)=(1,1)=2\text{cm}$ and $(0,0)=0\text{cm}$ per loci.

 30

```
1 @bind num_groups Slider(1:30,30,true)
```

 6

```
1 @bind number_loci Slider(0:15,6,true)
```

 3

```
1 @bind min_num_hetero Slider(0:15,3,true)
```

Different number of dominant loci can be selected

 2

```
1 @bind number_of_dom Slider(0:15,2,true)
```

The effect of A1 is +1cm. An A1A1 individual would have +2cm.

 1

```
1 @bind effect_A1 Slider(-3:3,1,true)
```

 0

```
1 @bind effect_A2 Slider(-3:3,0,true)
```

The effect of dominance can be toggle to simulate partial dominance, total dominance, underdominance and overdominance. d is the effect over the mid-effect value, $\text{mean}(A1,A2)$ $d=0$ (additivity), $d=1$ (total dominance)

```
1 md""
2 The effect of dominance can be toggle to simulate partial dominance, total dominance,
  underdominance and overdominance. d is the effect over the mid-effect value,
  mean(A1,A2)
3 d=0 (additivity), d=1 (total dominance)
4 ""
```

 0

```
1 @bind effect_d Slider(-2:2,0,true)
```

The baseline is 10cm.

```
1 md""
2 The baseline is 10cm.
3 ""
```

 10

```
1 @bind baseline Slider(-20:20,10,true)
```

In the activity we model a uniform distribution by rolling a die, however it may be useful to explore a normal distribution of environmental effects.

Uniform ▾

```
1 @bind env_type Select(["Uniform","Normal"])
```

env_effect is a multiplier of the environmental range.

 1

```
1 @bind env_effect Slider(-3:3,1,true)
```

env_sd is a parameter to model the normal distribution variance.

```
1 md""
2 env_sd is a parameter to model the normal distribution variance.
3 ""
```

 2.0

```
1 @bind env_sd Slider(-3.0:3.0,2.0,true)
```

Epistasis can also be modeled by adding a value if a minimum of loci (epistasis_level) has a value of A1 per loci equal or greater than N.

```
1 md""
2 Epistasis can also be modeled by adding a value if a minimum of loci
  (epistasis_level) has a value of A1 per loci equal or greater than N.
3
4 ""
```

false ▾

```
1 @bind epistasis Select([false,true])
```

 3.0

```
1 @bind epistasis_level Slider(0.0:6.0,3.0,true)
```

 10.0

```
1 @bind epistasis_value Slider(-10.0:10.0,10.0,true)
```

7

get_offspring (generic function with 1 method)



RObject{StrSxp}
[1] "p2.png"

Saving 7 x 7 in image

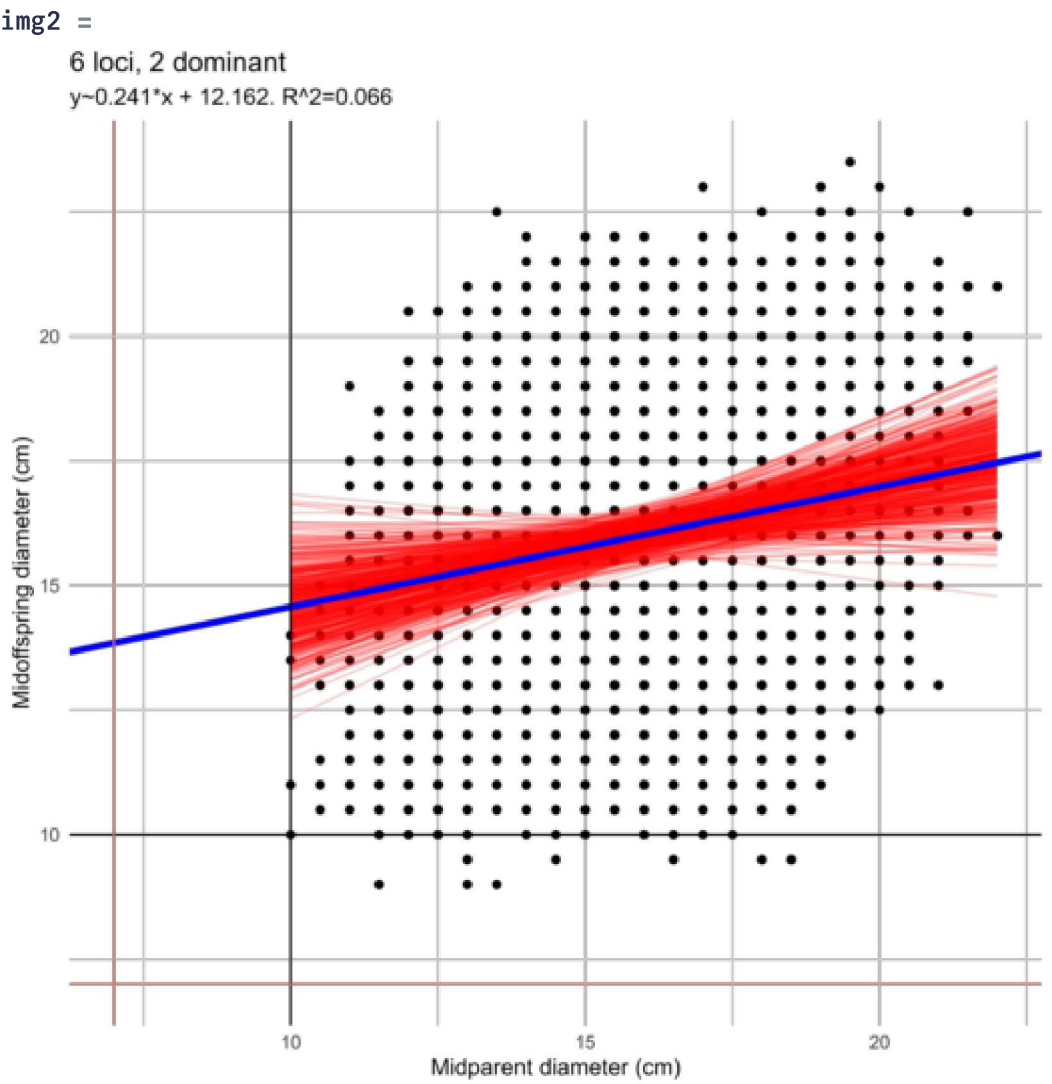
```
RObject{StrSxp}  
[1] "p4.png"
```

```
1 begin  
2  
3  
4     "Heritability less than 0=" * string(sum(slope_vec.<=0)/length(slope_vec))  
5     "Heritability less than 0.45=" * string(sum(slope_vec .<= 0.45) /  
length(slope_vec))  
6     "Average Heritability" * string(mean(slope_vec))  
7  
8     df2= DataFrame(slope=slope_vec,intercept=intercept_vec)  
9  
10    @rput df2  
11  
12    R""  
13    scatter <- ggplot(df2,aes(x=slope,y=intercept))+  
14        theme_minimal()+  
15        geom_hex()  
16  
17    hist_bottom = ggplot()+  
18    geom_histogram(data=df2,aes(slope))+  
19    theme_minimal()  
20    hist_right = ggplot()+  
21    geom_histogram(data=df2,aes(intercept))+  
22    coord_flip()+  
23    theme_minimal()  
24  
25    empty <- ggplot() +  
26        geom_point(aes(1, 1), colour="white") +  
27        theme(axis.ticks = element_blank(),  
28            panel.background = element_blank(),  
29            axis.text.x = element_blank(),  
30            axis.text.y = element_blank(),  
31            axis.title.x = element_blank(),  
32            axis.title.y = element_blank())  
33  
34    p3 <- grid.arrange(scatter,  
35        hist_right,  
36        hist_bottom,  
37        empty,  
38        ncol=2, nrow=2, widths=c(4, 1), heights=c(4, 1))  
39  
40    ggsave(p3,file="p3.png",dpi=600)  
41  
42    p4 <- grid.arrange(p2,p3,ncol=2, nrow=1,widths=c(6, 8),heights=4)  
43    ggsave(p4,file="p4.png",dpi=600)  
44    ""  
45  
46 end
```

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.  
Saving 7 x 7 in image  
Saving 7 x 7 in image
```

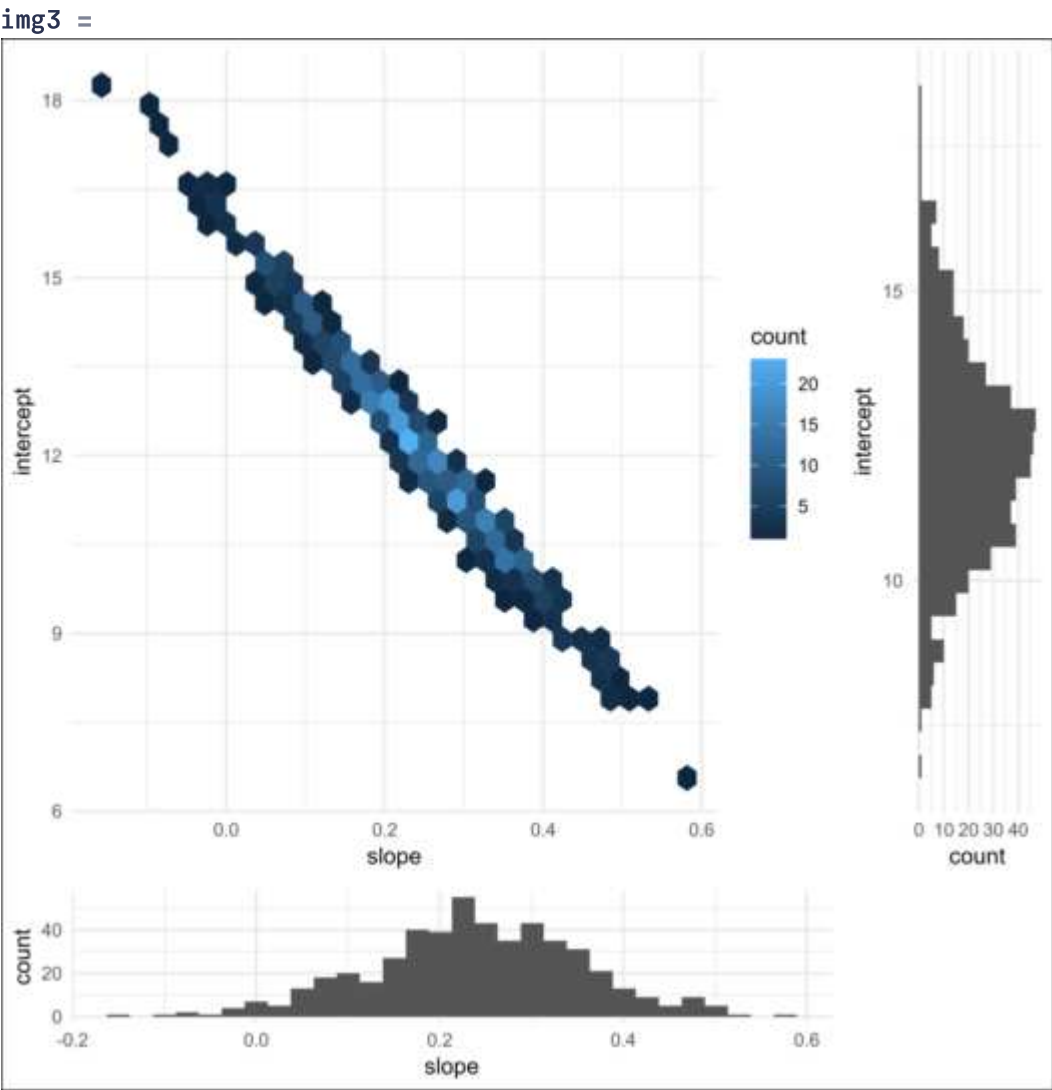


The slope is the heritability for all simulations (blue line). Each red line represent and individual simulation.



```
1 img2 = load("p2.png")
```

This plot shows the slopes and intercepts of all simulations.



```
1 img3 = load("p3.png")
```