Exercise 1 – experimental design

Based on: https://scienceparkstudygroup.github.io/rna-seq-lesson/02-experimental-design-considerations/index.html

Each question is 16 point (6x16= 96), 4 point given for submitting.

Case 1:

A scientist has to measure the effect of a plant pathogen called "*Designus malatesta*" on two plant genotypes: one resistant to the pathogen (labelled "*R*" for resistant) and one susceptible (labelled "*S*").

A total of 24 individual plants are placed on two tables in a greenhouse compartment. Each table can accommodate 12 plants. The greenhouse upper right side is pointing north, close to the location of the entrance door. Finally, there is an electrical board to power lamps and other high voltage installation at the lower left side opposite of the door.

1. Which **experimental factors** are controlled by the scientist?

Answer: There are two controlled factors:

A. The **genotype factor** with two levels: *R* and *S*
B. The **infection factor** also with two levels: uninfected (control) and infected.

2. Define the **conditions** in the experiment, and state how many replicates per condition can be used by the scientist (using this greenhouse)?

Answer: There are 4 conditions:

A. The resistant plant *R* uninfected
B. The susceptible plant *S uninfected*
C. The resistant plant *R* infected
D. The susceptible plant *S infected*

There can be up to 6 replicates per condition.

3. Can you identify another potential source of unwanted variation due to the practical setup of the experiment?

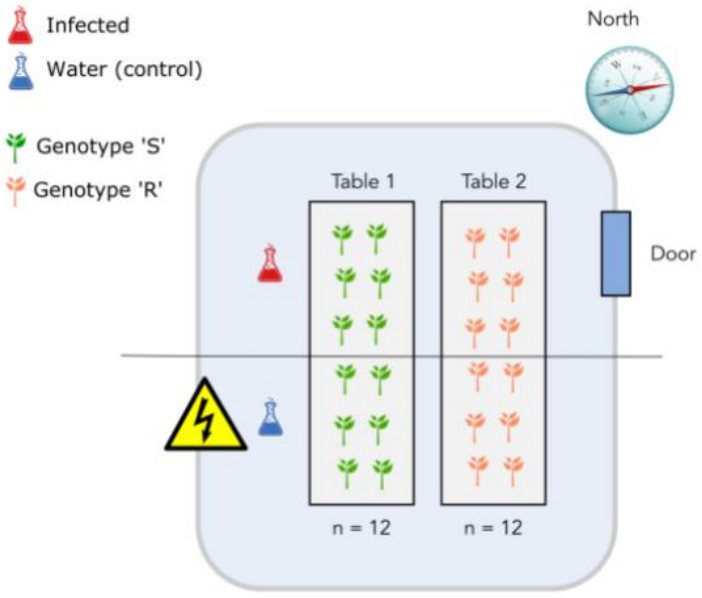
Answer: Another experimental factor that is undesirable can be called "temperature" and will cause table 1 and table 2 to be different.

The greenhouse compartment door is on the North-East side so potentially a source of cold air. On the contrary, there is an electrical board on the lower South-West side of the compartment, a potential source of warm air. Therefore, there might be a heterogeneity in the local temperature that plants are experiencing on table 1 and 2.

4. The scientist performed the pathogen inoculation using infected water on one part of the greenhouse and plain (control) water on the other part – described by a line in the drawing below.

Can you identify a problem with the suggested experiment setup?

How would you correct it?

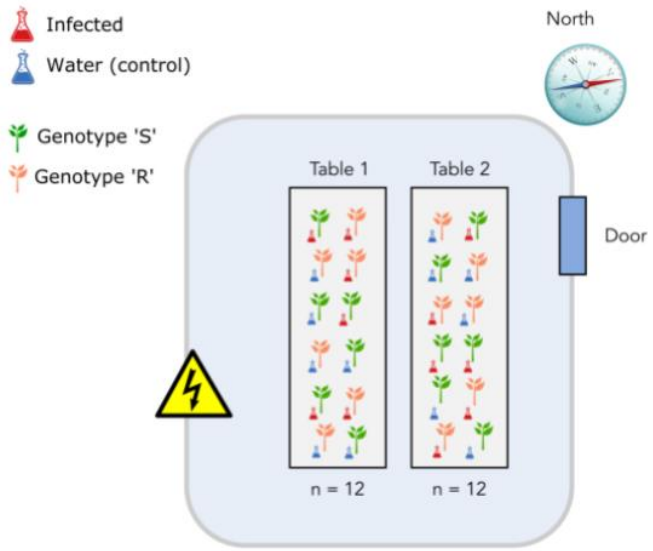


Answer: The major issue is that it is not possible to distinguish the "table effect" (undesirable) from the "genotype" effect since all *S* genotypes are on table 1 while all *R* genotypes are on the table 2.

The best experimental setup would randomize both the positions of the plant genotypes and the treatment on each table. In this design, each table will accommodate a complete experimental setup. This means that each table will exhibit:

- Plant genotypes S and R in equal numbers.
- Infection or control (water) conditions in equal numbers.

An example of the design:

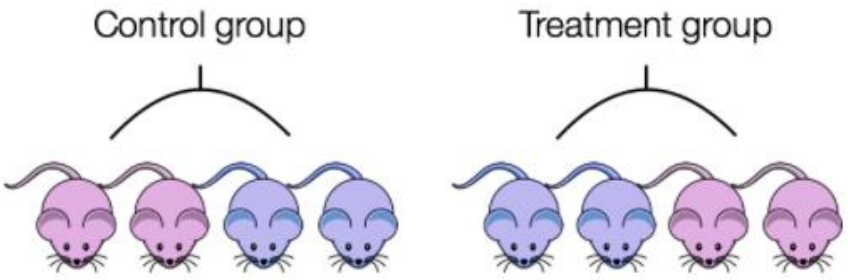


Case 2:

A scientist wants to measure the effect of a drug treatment on mice. There are 8 mice, 4 female and 4 male.

5. Suggest how to divide the mice between the conditions.

Answer: We want to ensure animals in each condition are all the same sex, age, litter, and batch, if possible. If not possible (as in this example), then ensure to split the animals equally between conditions.



Case 3:

Due to equipment limitations, you can perform simultaneously up to 8 RNA isolations. You have the following samples (image without batch information).

6. Suggest how to perform RNA isolation (divide the samples between the batches). Fill the batch column in the table.

Answer: We want to split replicates of the different sample groups equally across batches, and minimize number of batches.

Sample	Replicate	Condition	Batch
sample1	1	Control	1
sample2	2	Control	1
sample3	3	Control	2
sample4	4	Control	2
sample5	1	Treatment A	1
sample6	2	Treatment A	1
sample7	3	Treatment A	2
sample8	4	Treatment A	2
sample9	1	Treatment B	1
sample10	2	Treatment B	1
sample11	3	Treatment B	2
sample12	4	Treatment B	2