Australian National University

Statistical Thinking in Biology Research
BIOL 8291

Terry Neeman, PhD AStat
Statistical Consulting Unit (SCU)
John Dedman Building, ANU

Let’s talk about statistics and its role in the sciences.

Tick all that apply:

Statistics is important in:

- ❑ Biological sciences
- ❑ Chemistry
- ❑ Physics
- ❑ Mathematics

What are sources of variation in biological experiments?

Tick all that apply:

- ❑ Environment (temperature, light, etc)
- ❑ Biological samples
- ❑ Technicians
- ❑ Machines, chemicals, proteins
- ❑ Experimental factors of interest

Statistical thinking in biological research:

Thinking systematically about all aspects of the experiment

How will this (experimental) factor impact my outcome?

How will these factors jointly impact my outcome?

What are other sources of variation that affect ~~the~~ outcome?

How can I design my experiment efficiently?

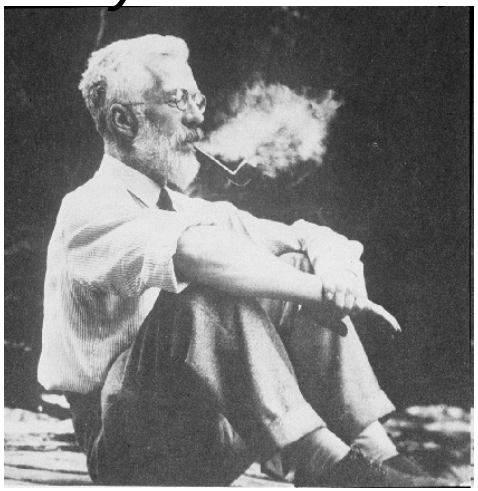
The Anatomy of an Experiment

What is your research question?

What is the main outcome of interest?

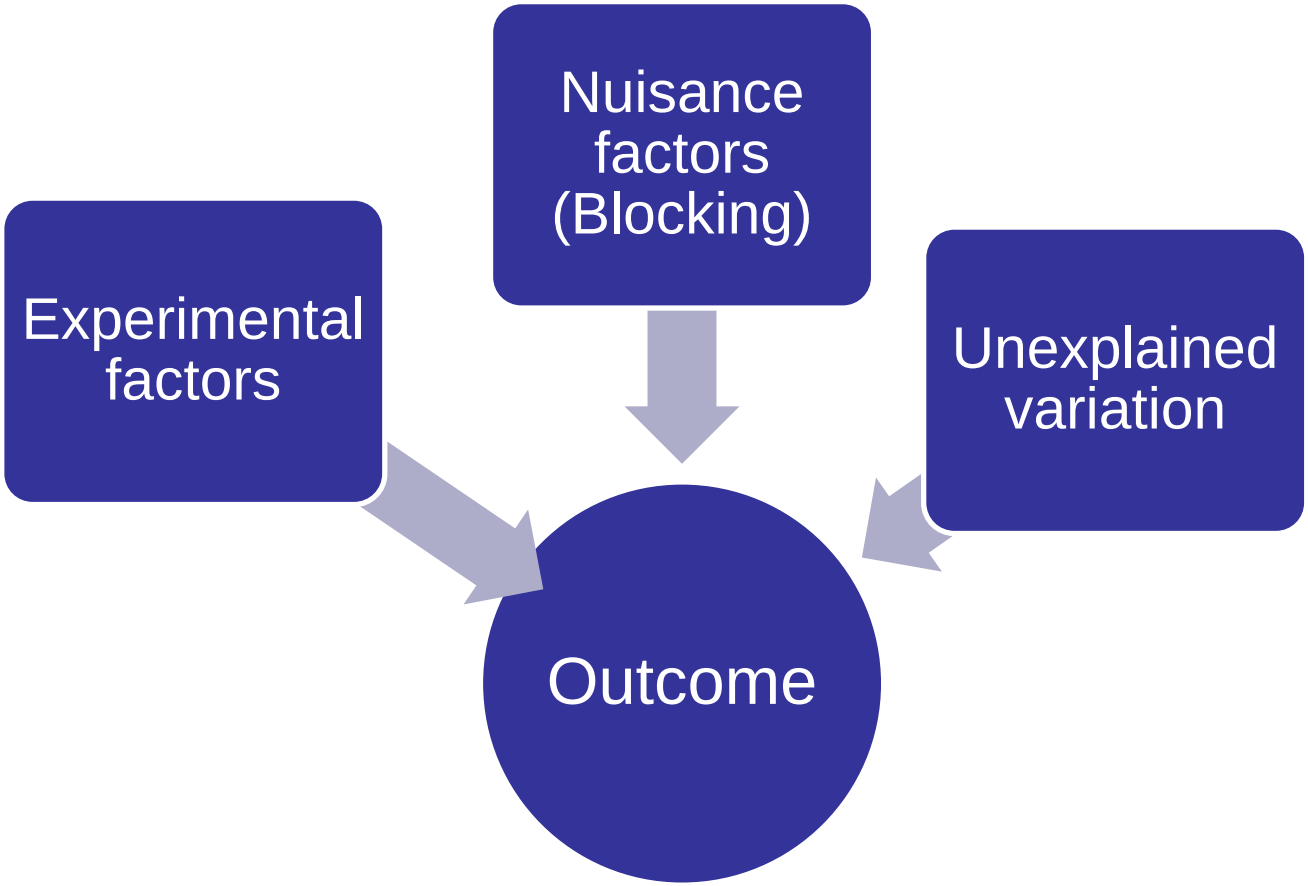
What the experimental factors?

What are other factors in the experiment that may influence your outcome?



R.A. Fisher, Statistical Methods for Research Workers (1925)

Anatomy of an Experiment

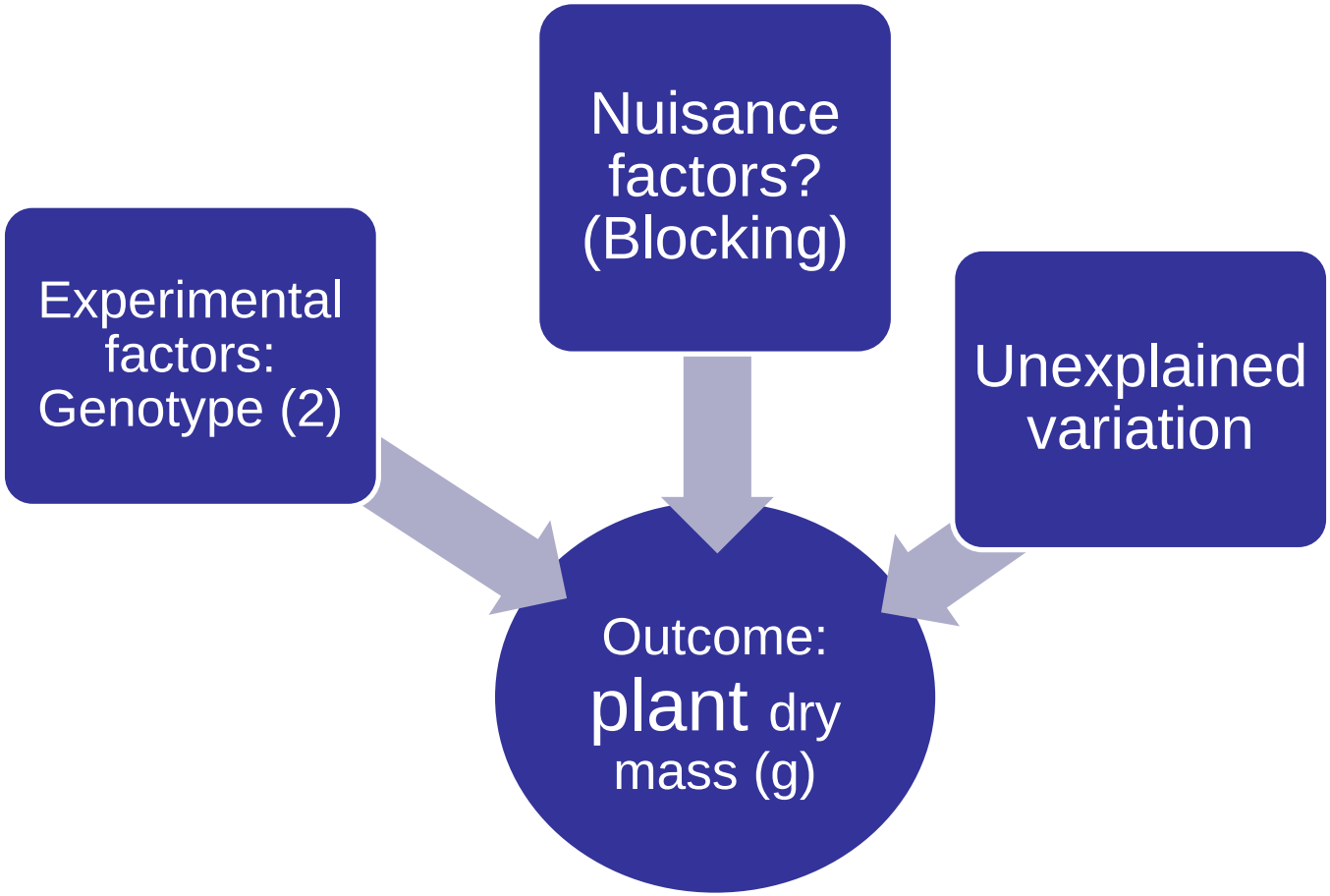


A Simple Experiment:

I have knocked out an protein in a tomato plant variety that I believe will result in a drought-resistant plant that produces excellent juicy tomatoes.

In my first experiment, I'll test if the drought-resistant seedlings grow at a similar rate to the wild-type controls.

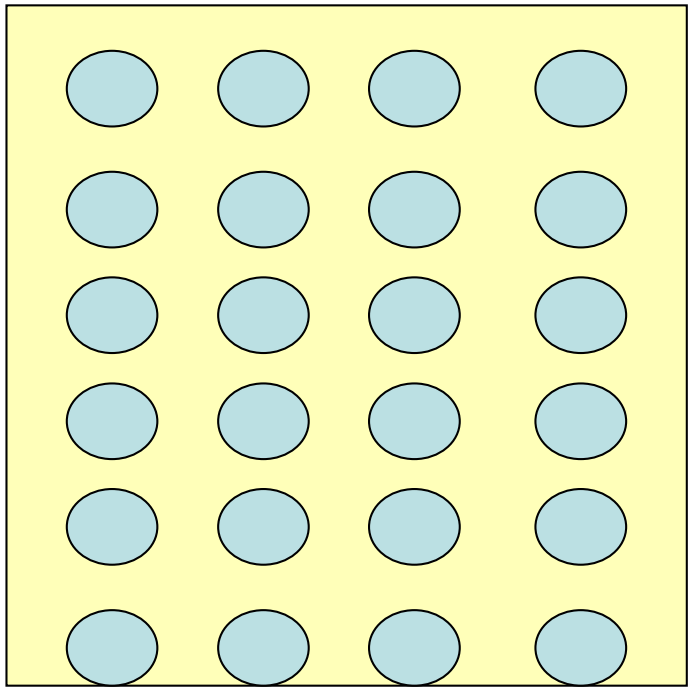
Anatomy of an Experiment



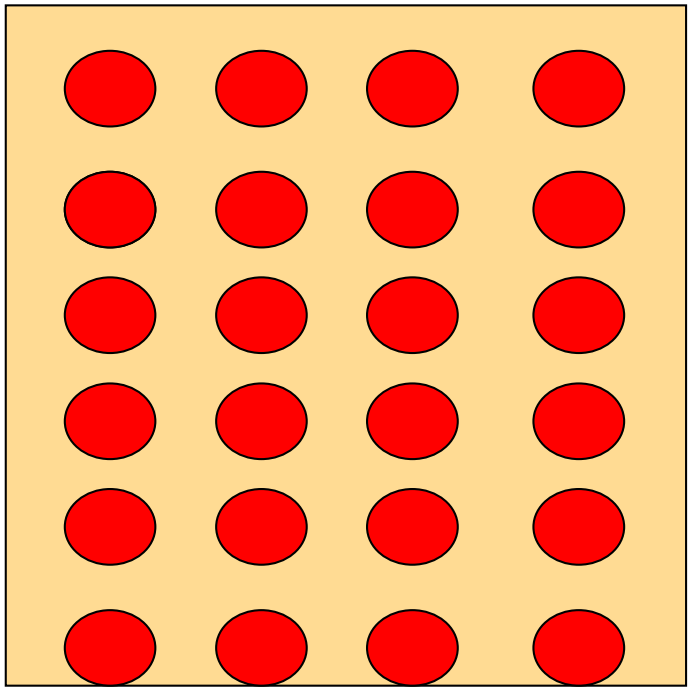
Comparison of 2-week growth of tomato plants:

(i) wild-type & (ii) drought-resistant

Top shelf: wild-type

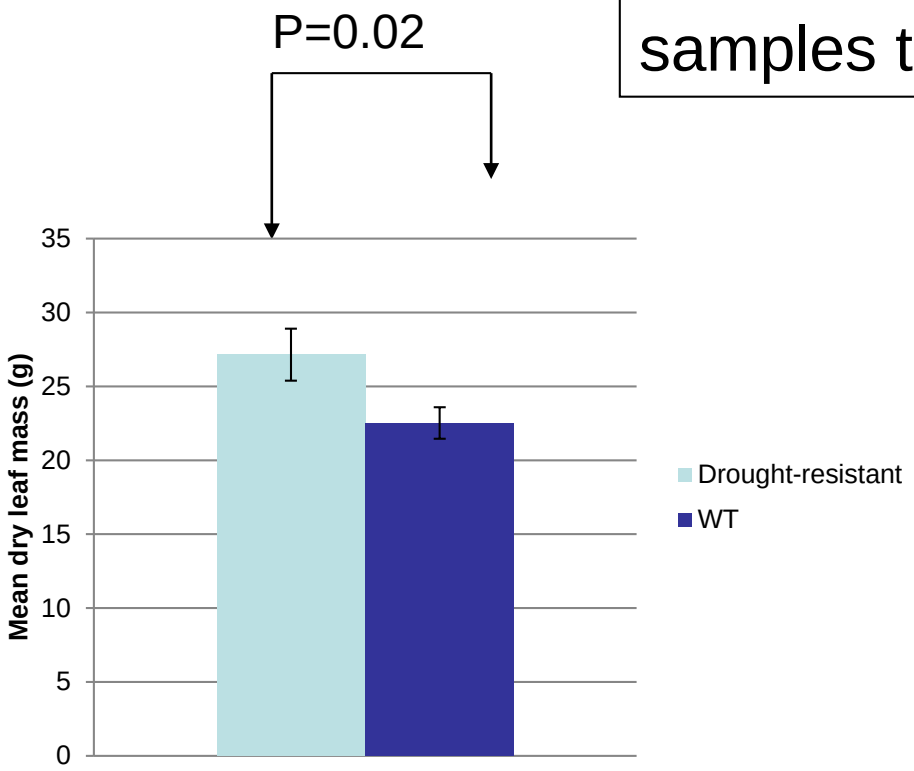


Bottom shelf: drought- resistant



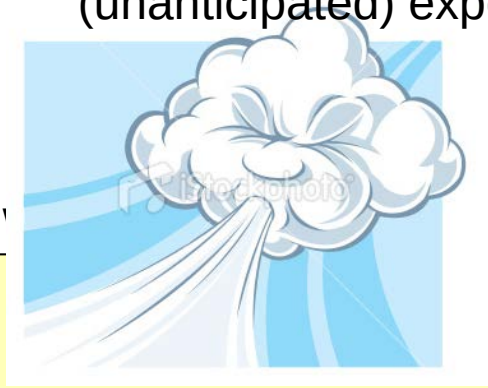
Results: plant dry mass (g) after 2 weeks

Statistical method: independent samples t-test, df = 46

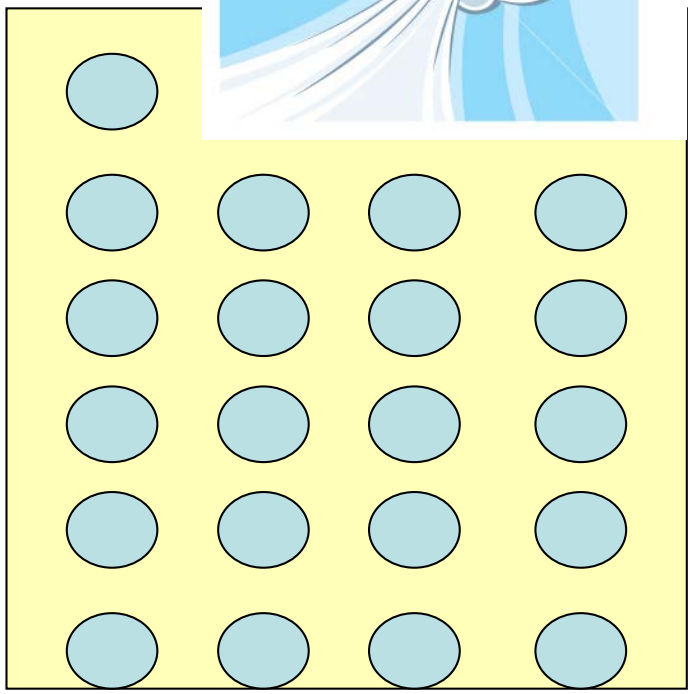


Comparison of 2-week growth of tomato plants:

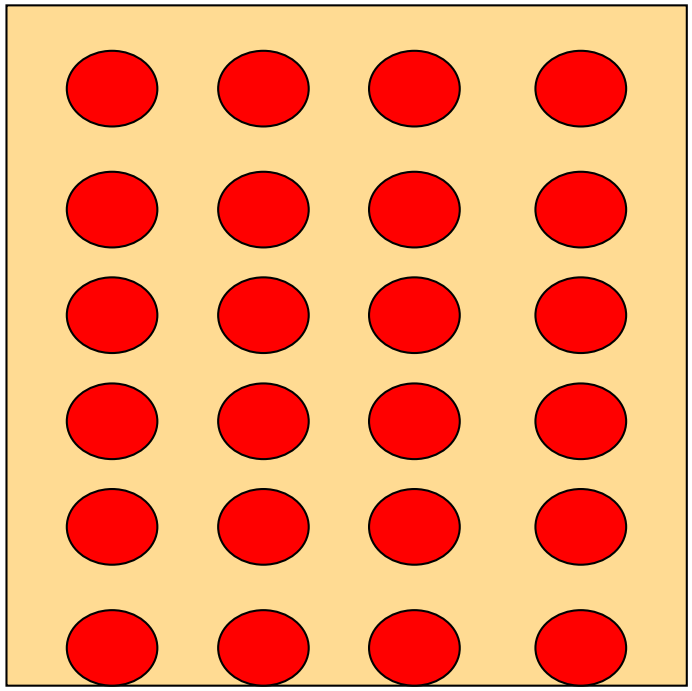
(unanticipated) experimental conditions



Top shelf:



Bottom shelf: drought- resistant



What are potential sources of (nuisance) variation in this experiment?

Tick all that apply:

- ❑ Shelf
- ❑ Position on shelf
- ❑ Technicians
- ❑ Machines, chemicals, proteins

How can we “control” for these sources of variation?

How do we “control” the variation due to nuisance factors?

- In the experiment above, shelf and genotype are confounded.
- Confounded: We cannot separate the “shelf effect” from the “genotype effect”.
- Orthogonal: If we can completely separate out the effect of nuisance factors from the effect of experimental factors of interest on the outcome, then we get the most information about how the experimental factors affect the outcome.

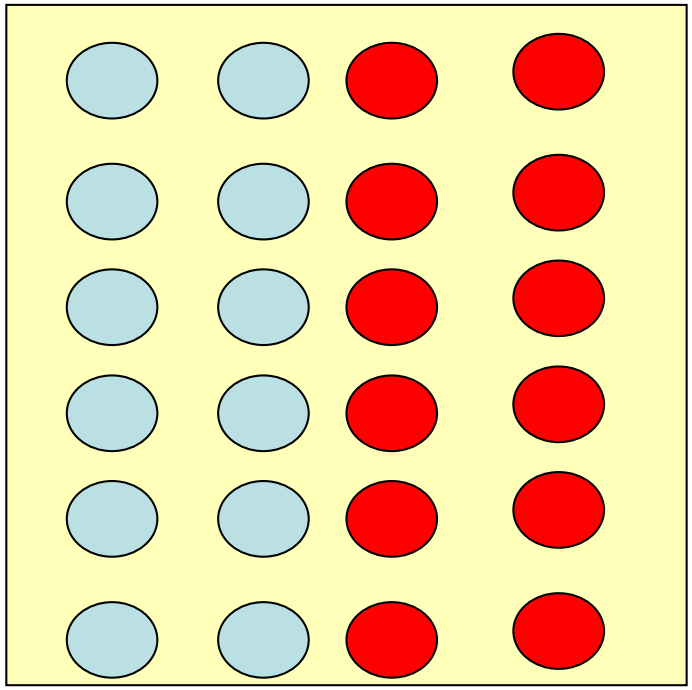
Key point: Good experimental design ensures that we get the most information about how the experimental factors influence the outcome.

Let's improve our experimental design

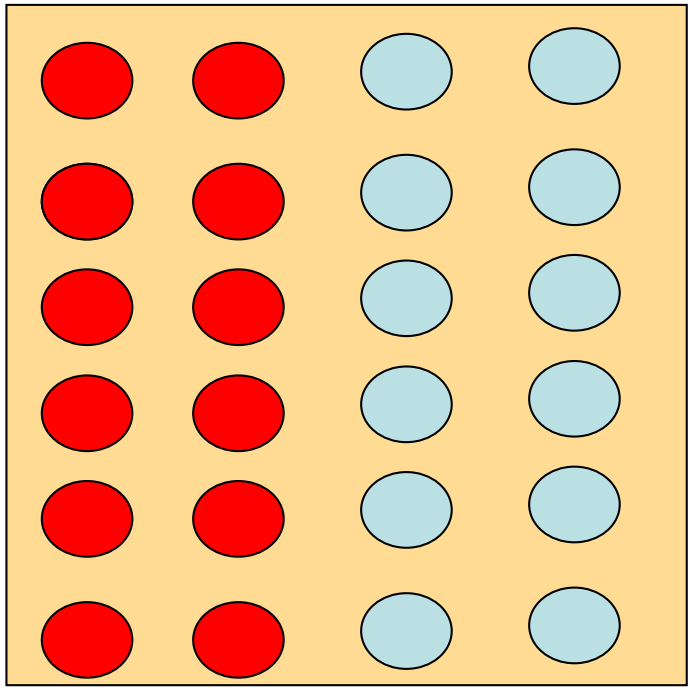
Comparison of 2-week growth of tomato plants:

Proposed design

Top shelf: wild-type/drought-resist



Bottom shelf: drought- resist/wild-type

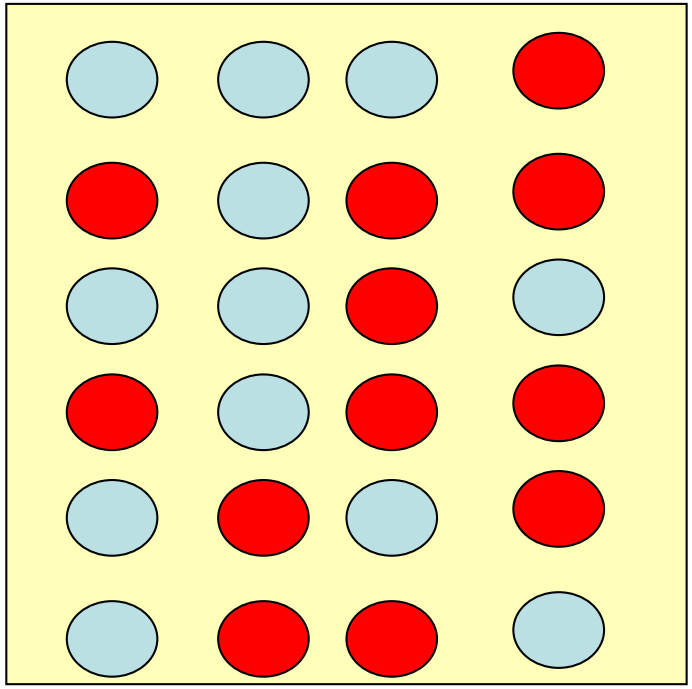


Let's improve our experimental design

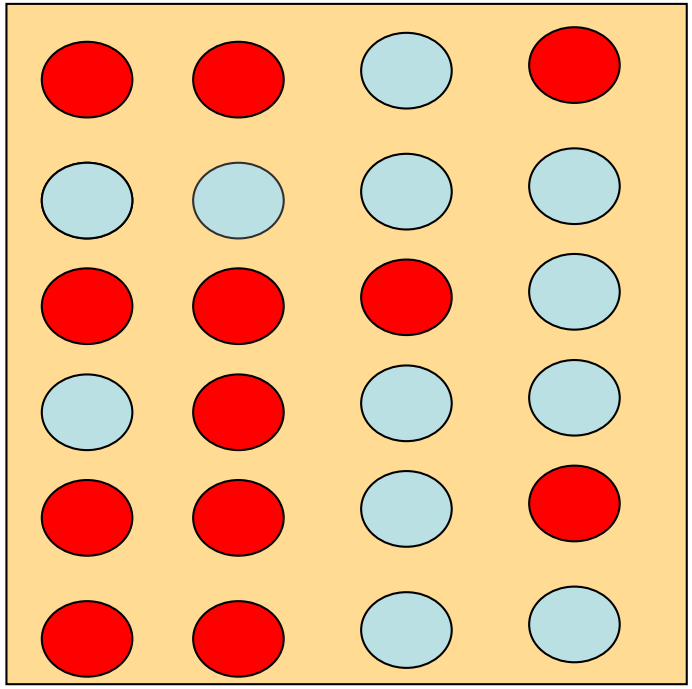
Comparison of 2-week growth of tomato plants:

Complete randomised block design

Top shelf: wild-type/drought-resist



Bottom shelf: drought- resist/wild-type

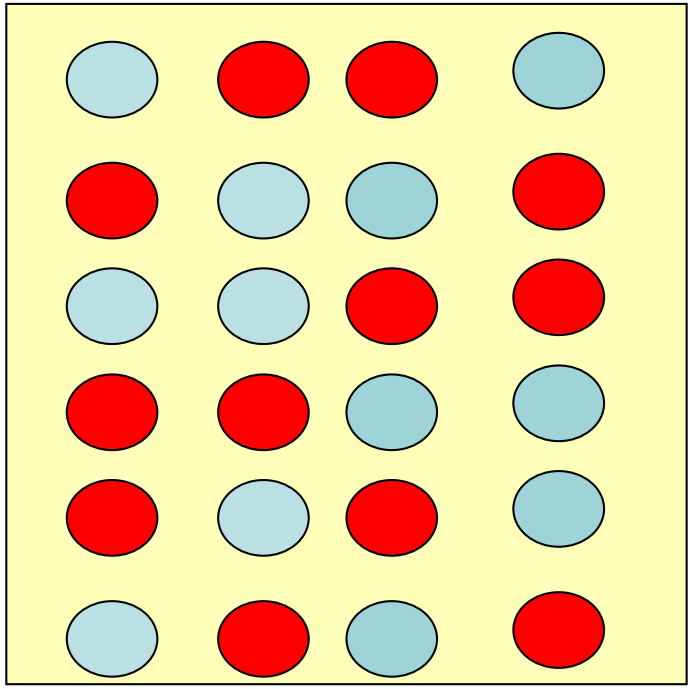


Let's improve our experimental design

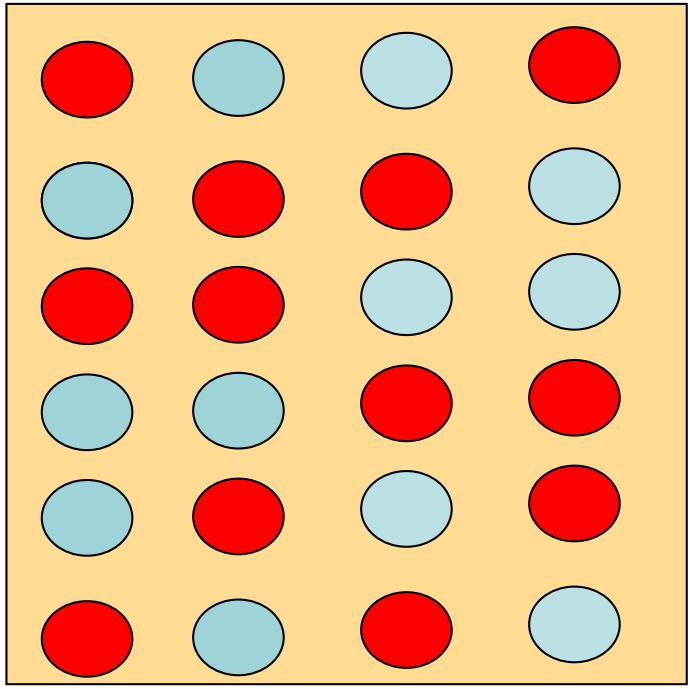
Comparison of 2-week growth of tomato plants:

Proposed design: Row-column design

Top shelf: wild-type/drought-resist



Bottom shelf: drought- resist/wild-type



Anatomy of an Experiment

- Clear research question that articulates:
  - Outcome measure
  - Experimental factors of interest
- Consideration of experimental conditions that may also impact the outcome measure: experimental design
- Incorporate experimental design into a statistical model that addresses research question

Key components of a statistical model of an experiment

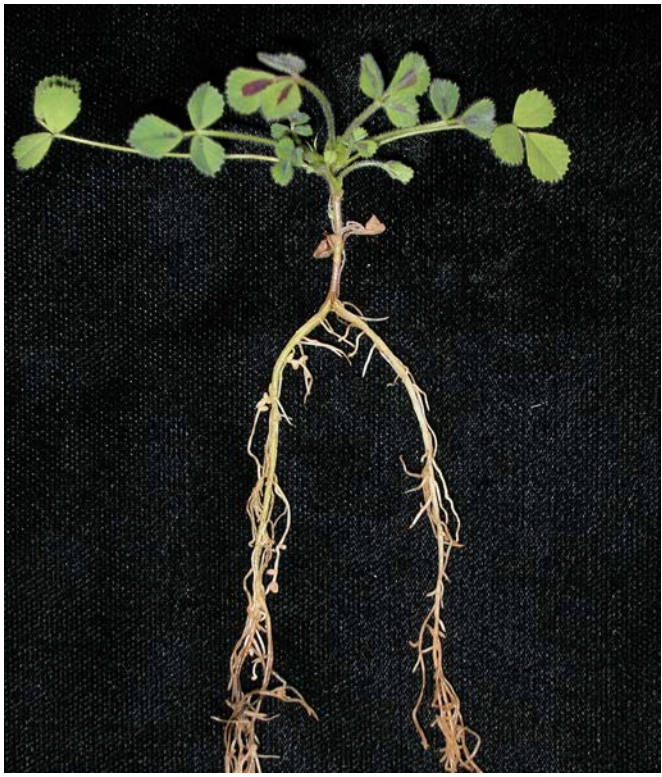
- Outcome measure
  - Response variable
  - Measure of interest
- Experimental factors
  - Conditions that can be manipulated
  - Conditions of interest (e.g. genotype, gender)
  - Main questions: do the conditions impact upon the outcome measure?
- Nuisance factors
  - Conditions (not of interest) that may impact upon the outcome measure
  - Sources of variation in the experiment that need to be controlled for
  - Clustering of experimental units

ALWAYS BEGIN WITH A RESEARCH QUESTION

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

•Research question for Experiment #1:

In legumes, are plastic responses to internal nitrogen altered by their rhizobial symbionts?



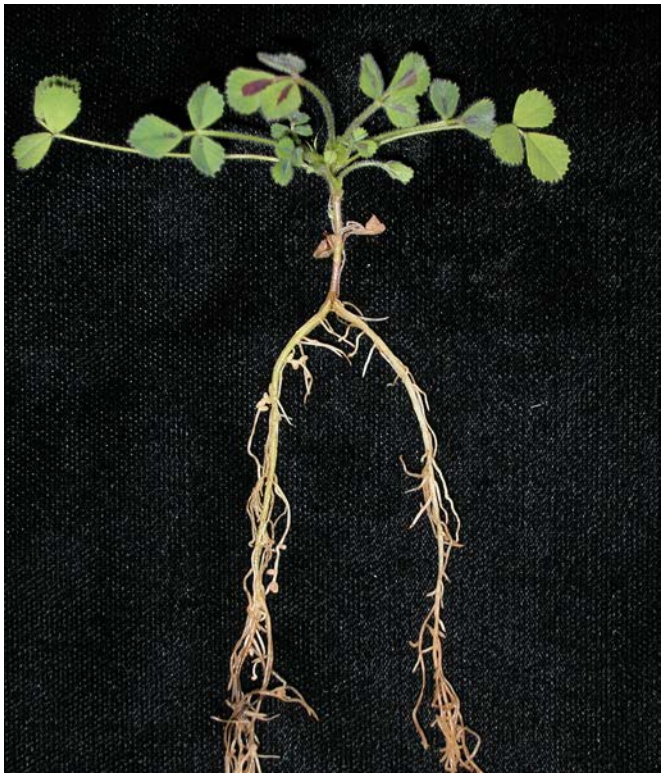
The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

- What were the outcome measures?

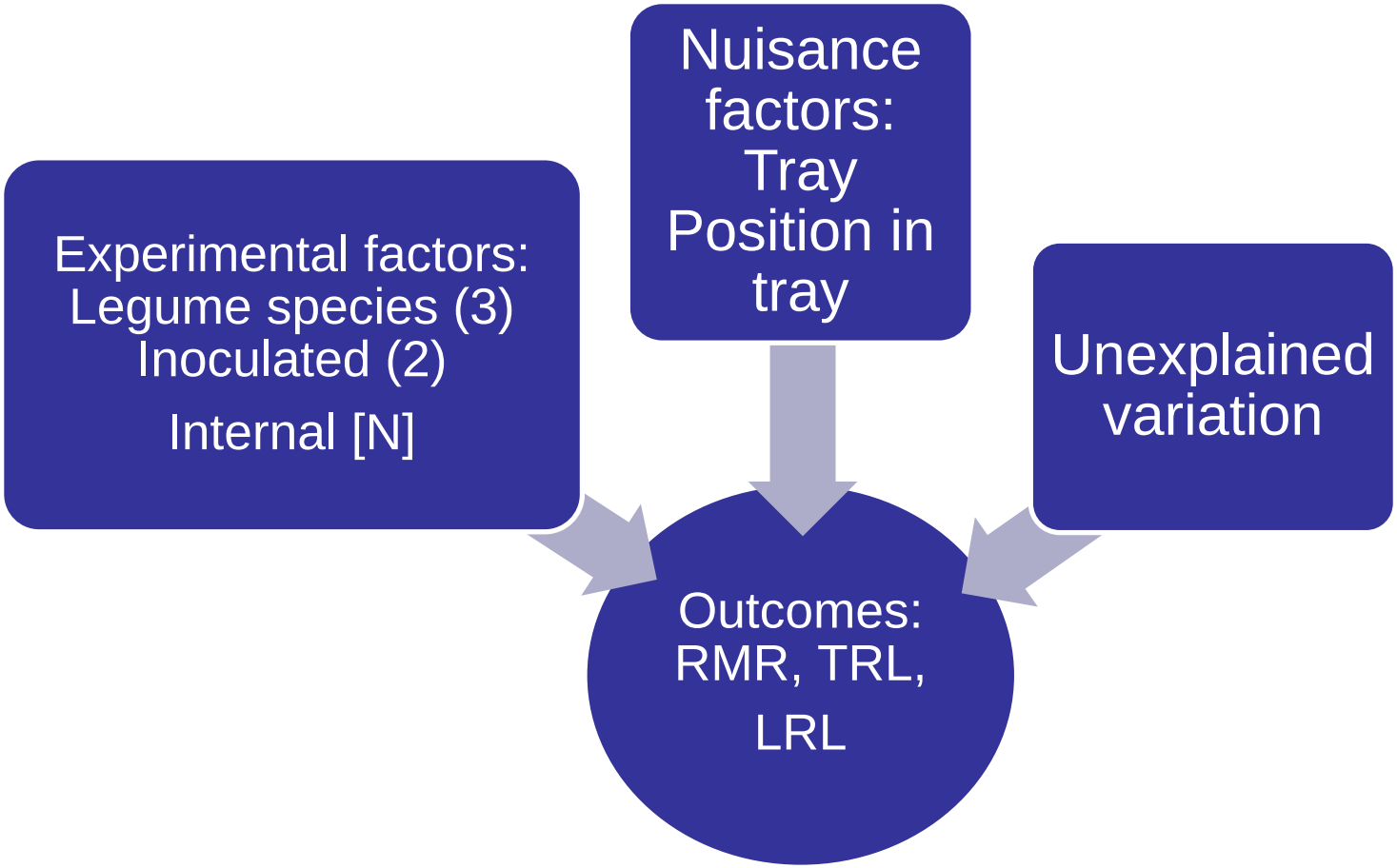
Root mass ratio (RMR),total root length (TLR), lateral root length (LRL)

- What were the experimental factors?

Species(3), Inoculated (2), nitrate (4)



Anatomy of this experiment



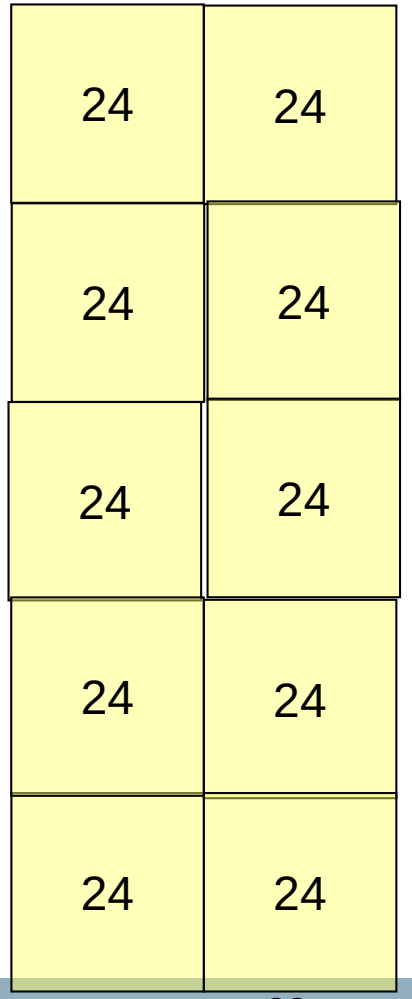
The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

- How many “conditions” and how many replicates per condition:

3 (species) x 2 (inoculate status) x 4 (nitrate levels) = 24 conditions

10 replicates/condition = 240 plants

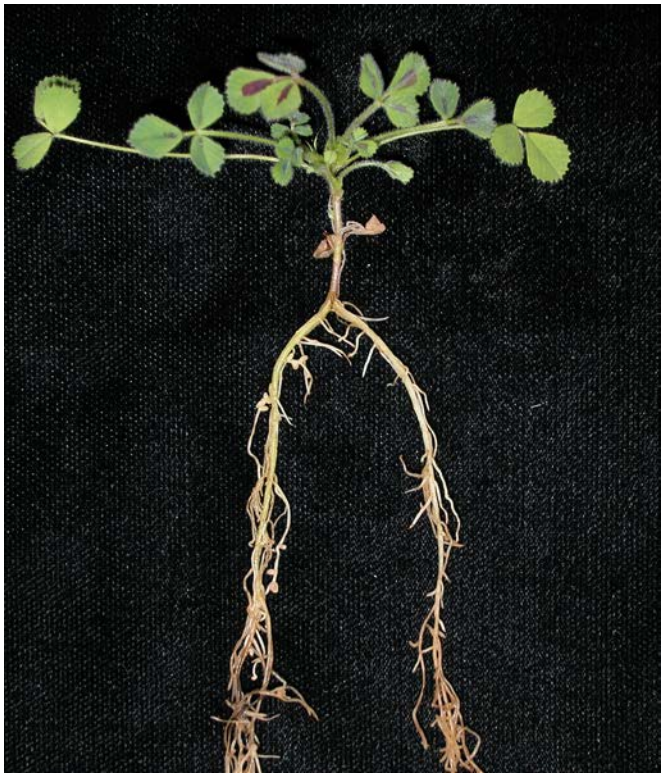
Can place 24 plants/tray



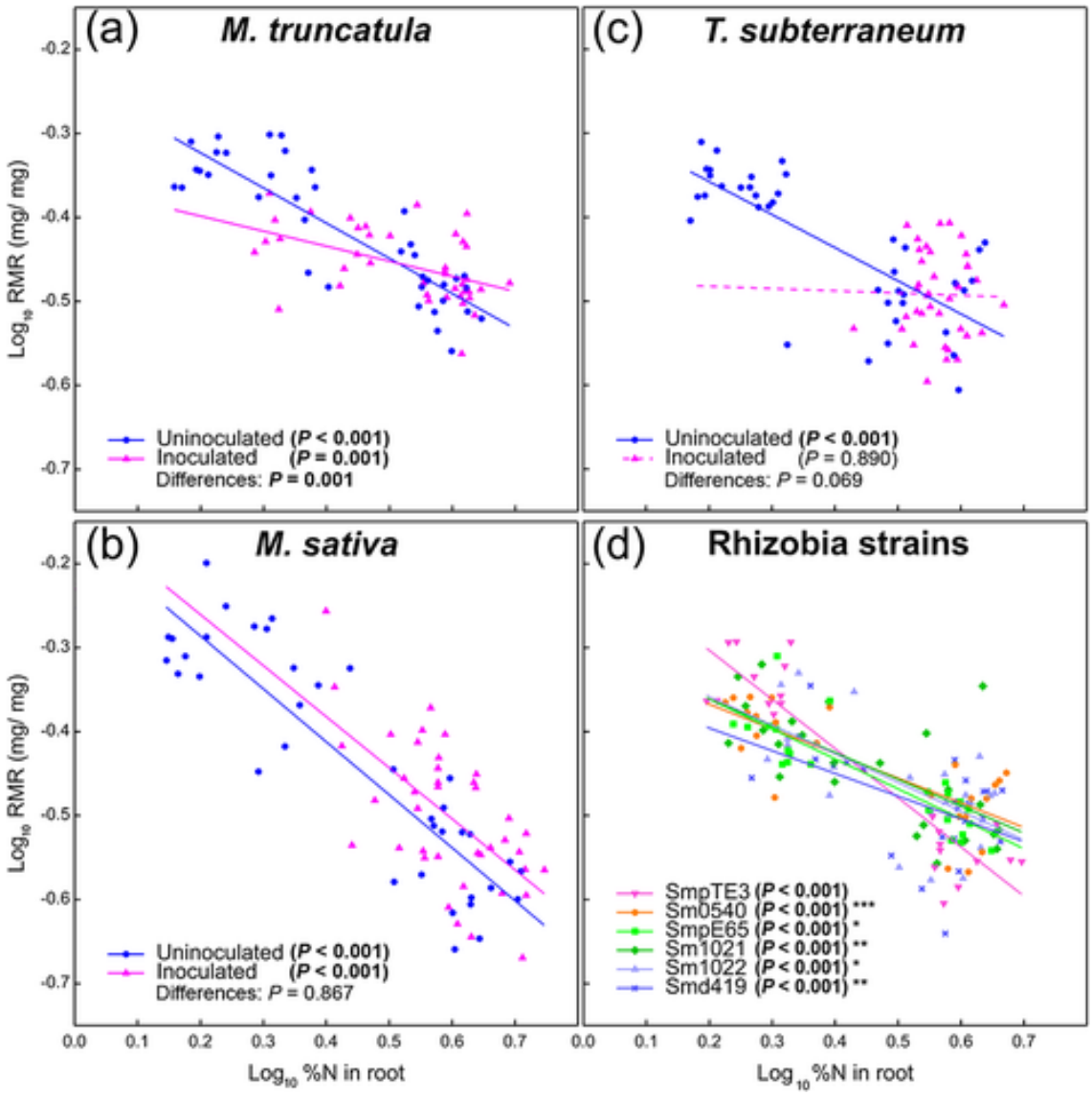
The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

•Experimental design:

“An efficient row-column design for a three-factor experiment”



The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation

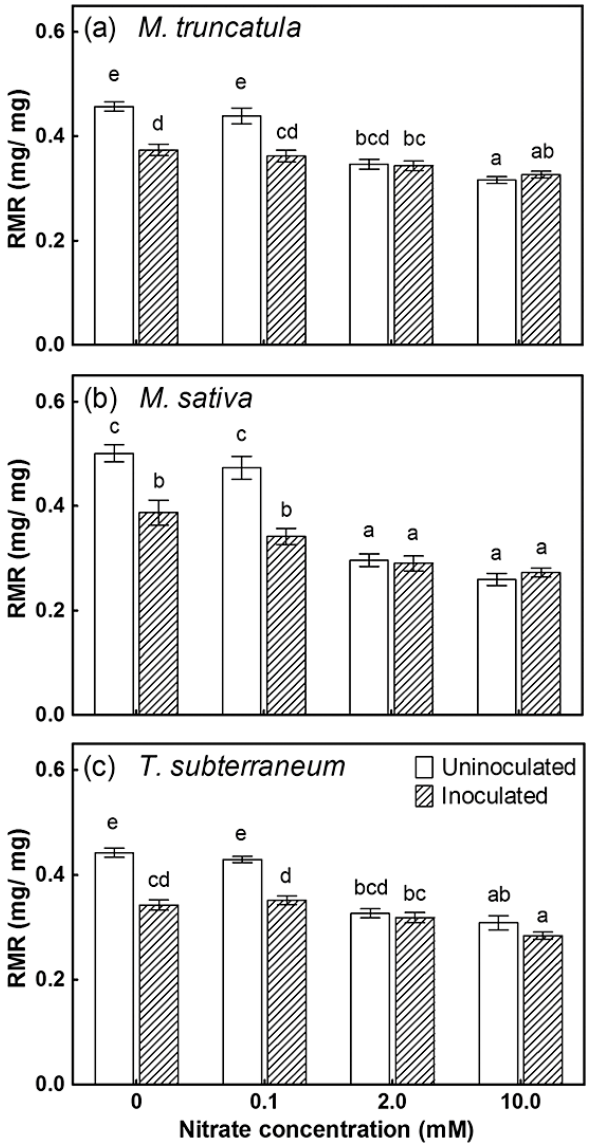


Interpret Figure 3a-c with respect to the research question:

Are plastic responses to internal nitrogen altered by their rhizobial symbionts?

Plant, Cell & Environment, Volume: 39, Issue: 4, Pages: 883-896, First published: 02 November 2015, DOI: (10.1111/pce.12672)

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation

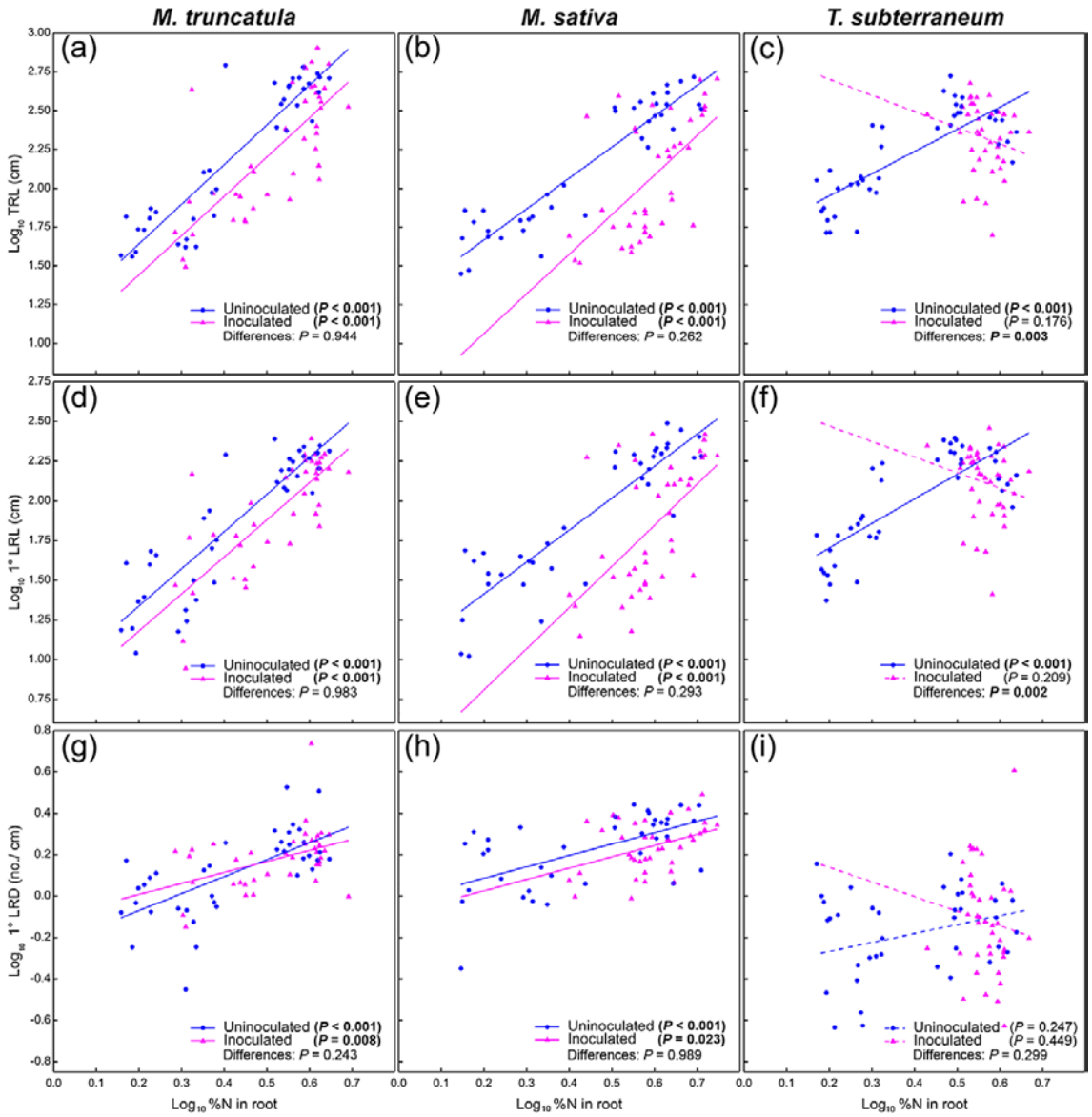


Interpret Figure 4 with respect to the research question:

Are plastic responses to nitrate treatment altered by their rhizobial symbionts?

Plant, Cell & Environment, Volume: 39, Issue: 4, Pages: 883-896, First published: 02 November 2015, DOI: (10.1111/pce.12672)

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation



Interpret Figure 5 with respect to the research question:

Are plastic responses to internal [N] altered by their rhizobial symbionts?

Plant, Cell & Environment, Volume: 39, Issue: 4, Pages: 883-896, First published: 02 November 2015, DOI: (10.1111/pce.12672)

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

•Research question for Experiment #2:

Do plastic responses to internal nitrogen vary by rhizobial strain in *M. truncatula*?

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

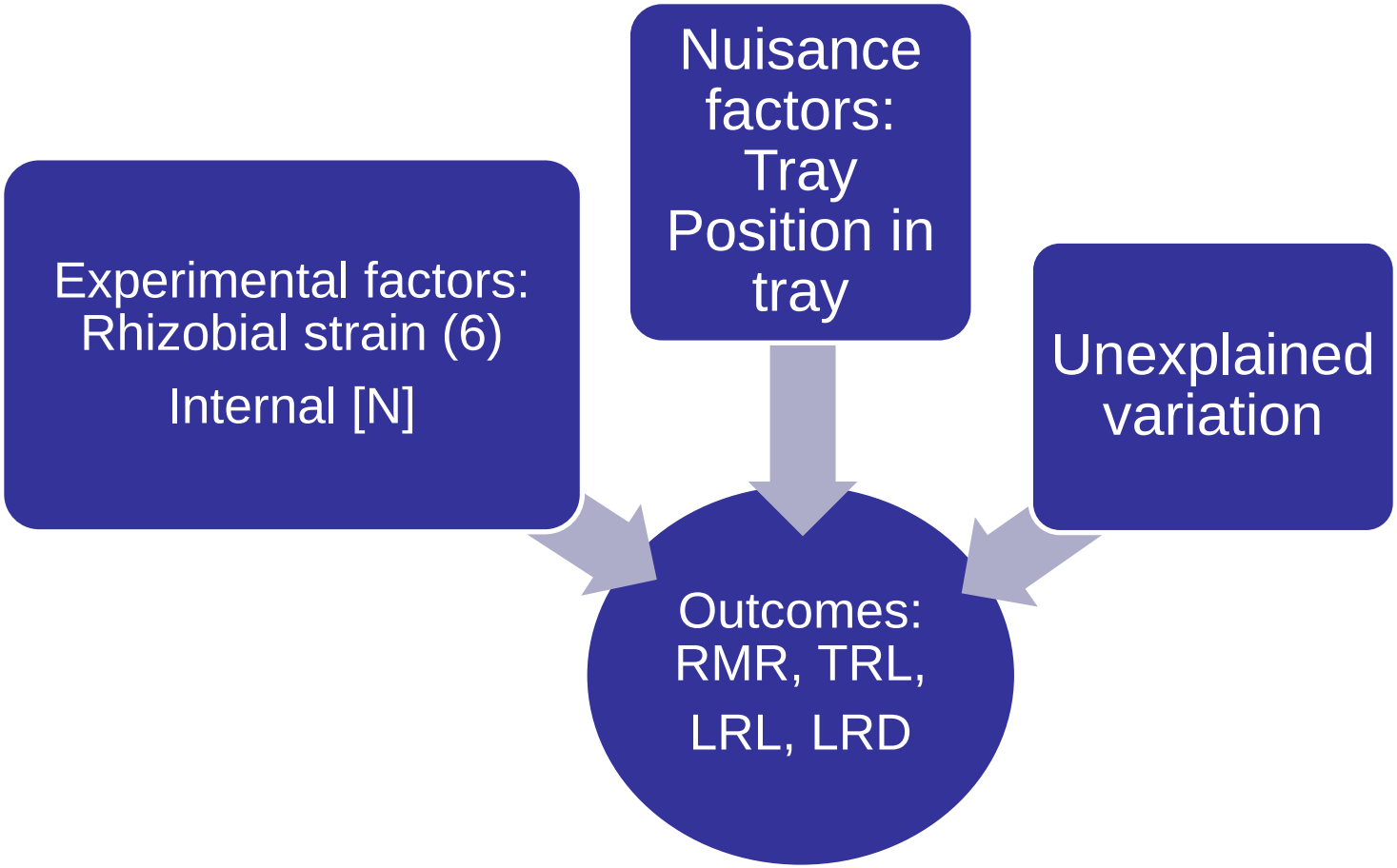
- What were the outcome measures?

Root mass ratio (RMR),total root length (TLR), lateral root length (LRL), lateral root density (LRD)

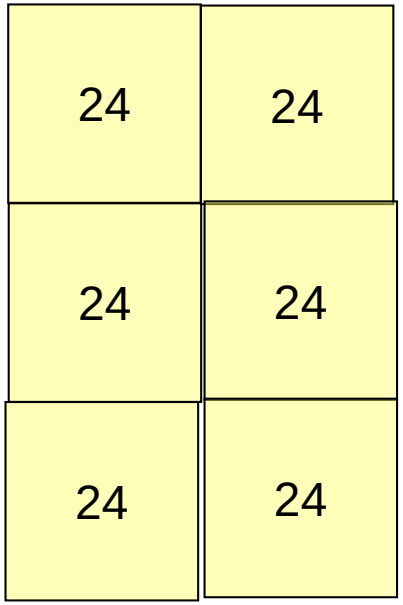
- What were the experimental factors?

Rhizobial strain (6), nitrate treatment (4)

Anatomy of this experiment



The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)



- How many “conditions” and how many replicates per condition:

6 (strains) x 4 (nitrate levels) = 24 conditions

6 replicates/condition = 144 plants

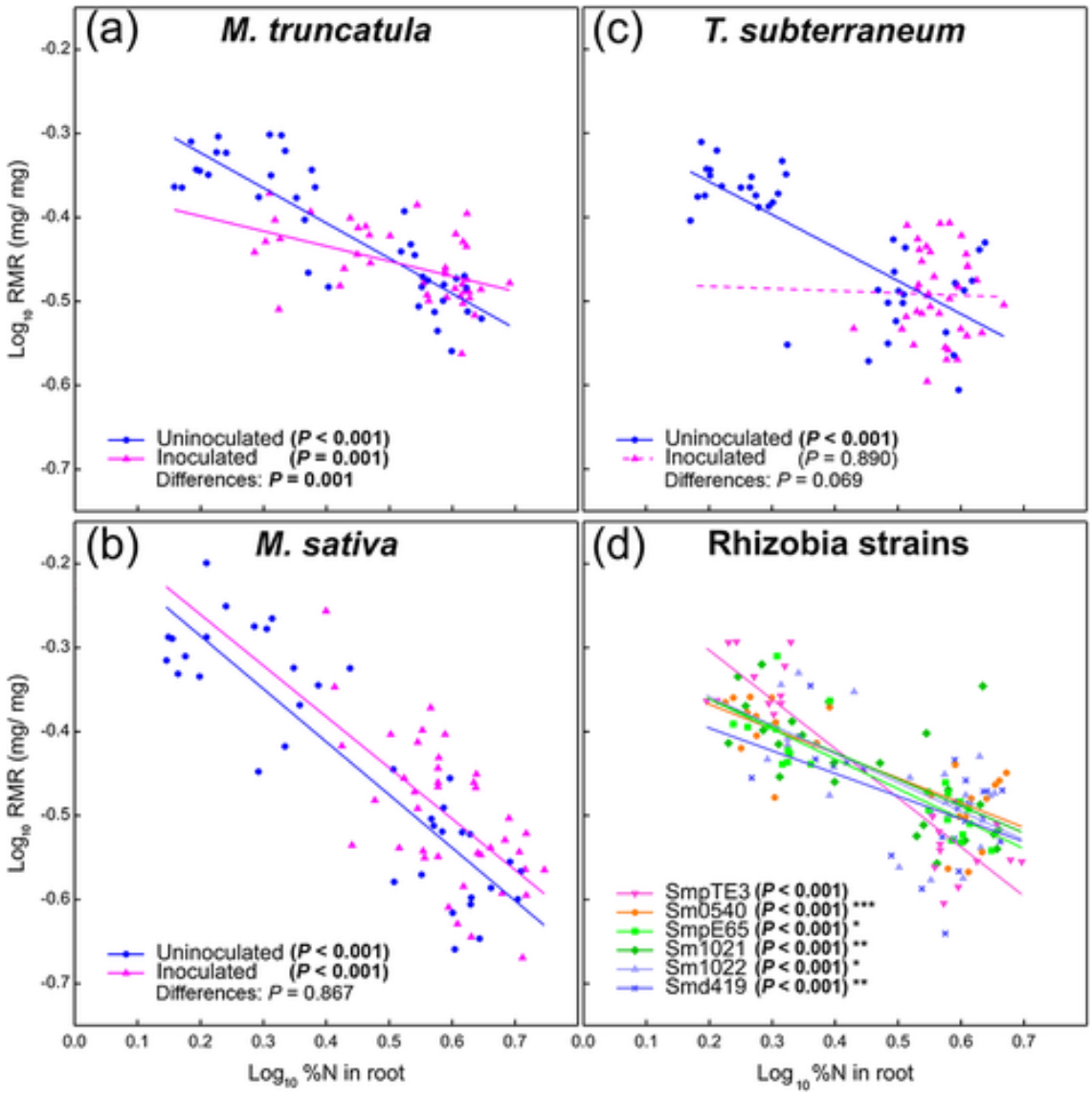
Can place 24 plants/tray

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation, Goh C. et al Plant, Cell and Environment (2016)

•Experimental design:

“A similar efficient resolvable row-column design for a two-factor experiment”

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation

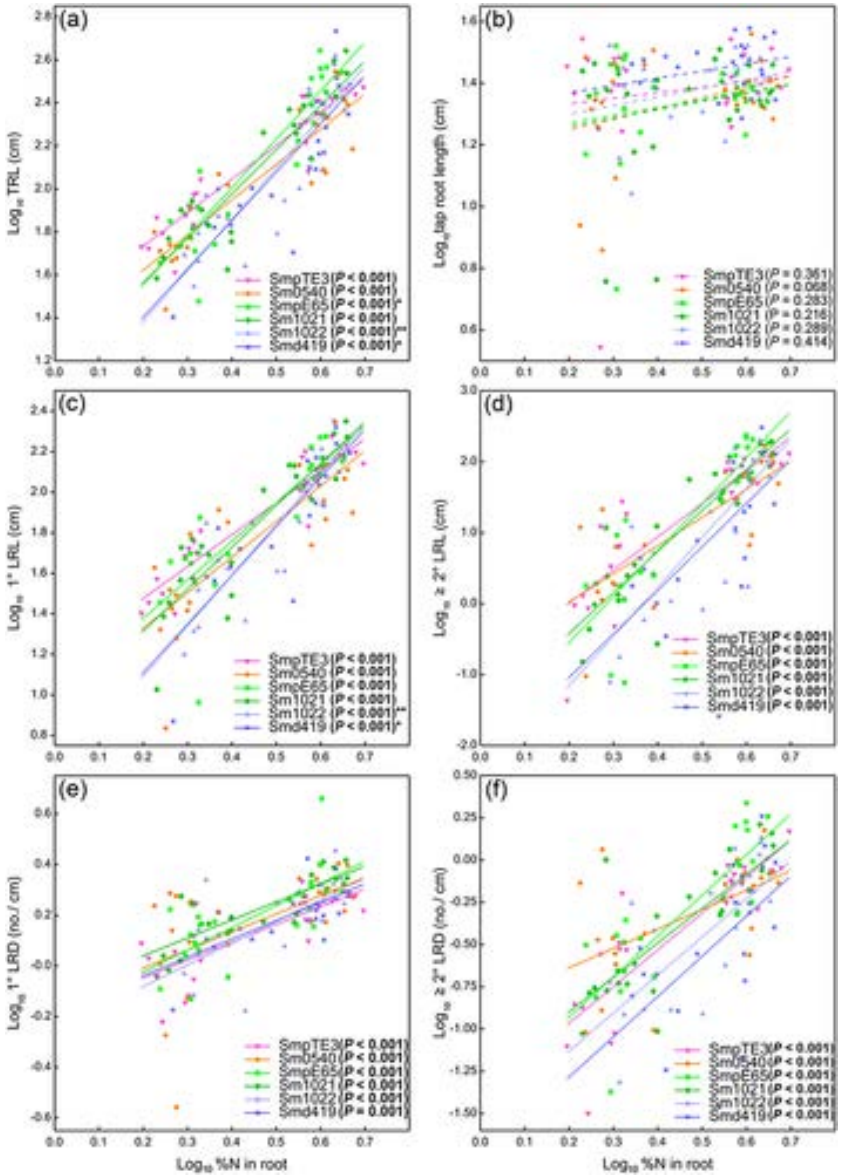


Interpret Figure 3d with respect to the research question:

Do plastic responses to internal nitrogen differ by strain in *M. truncatula*?

Plant, Cell & Environment, Volume: 39, Issue: 4, Pages: 883-896, First published: 02 November 2015, DOI: (10.1111/pce.12672)

The presence of nodules on legume root systems can alter phenotypic plasticity in response to internal nitrogen independent of nitrogen fixation



Interpret Figure 7 with respect to the research question:

Do plastic responses to internal nitrogen differ by strain in *M. truncatula*?

Plant, Cell & Environment, Volume: 39, Issue: 4, Pages: 883-896, First published: 02 November 2015, DOI: (10.1111/pce.12672)

Overall Summary

- Think systematically about all aspects of your experiment
- In your journal article, be able to identify (for each experiment):
  - The research question
  - Outcome(s) of interest
  - Experimental factors(s)
  - Nuisance factors(s)

- Statistical models are useful tools for estimating and assessing treatment effects, and answering research questions.

And for your research project next year…

- Think systematically about all aspects of your experiment
- Before you begin your experiment, identify:
  - The research question
  - Outcome(s) of interest
  - Experimental factors(s)
  - Nuisance factors(s)

- Talk it over with a statistician!

Statistical Resources available at ANU

- Biology Data Science Institute (BDSI)
- Statistical Consulting Unit (SCU)

Talk it over with a statistician!