Ciências ULisboa
Faculdade de Ciências da Universidade de Lisboa

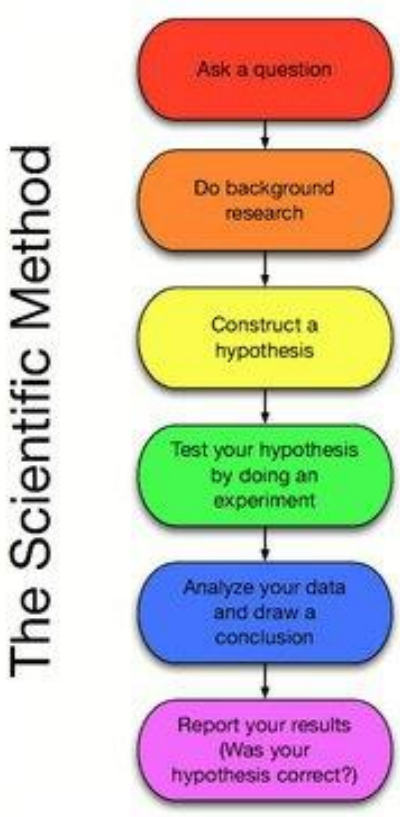
Experimental design and data analysis

Andreia Figueiredo, PhD

Faculty of Sciences University of Lisbon

aafigueiredo@fc.ul.pt

Scientific Method



1. Ask a question: What is the role of a protein “Z” in plant innate immunity?

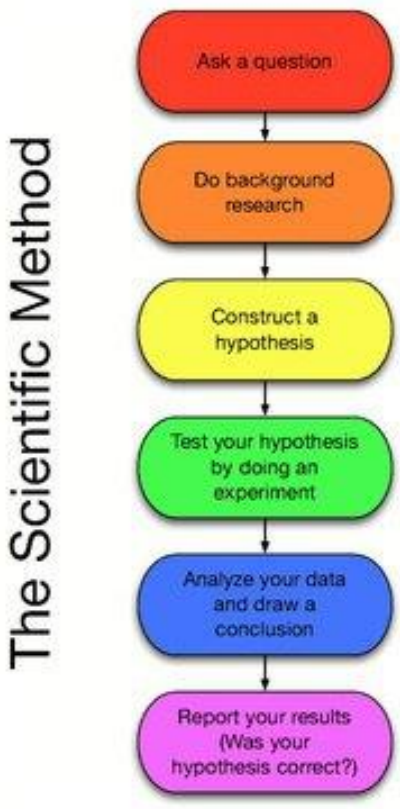
While asking your question you also need to decide an approach. The approach sets the theme of how you are going to address the question. Are you going for a top-down or bottom-up design?

(it is always advisable to break down your question into multiple small parts, and answer them individually as you move ahead)

2. Do background research: Search for research and review articles on PubMed, Google Scholar etc using keywords like “plant innate immunity”, “proteins involved immunity”, “immunity regulation in plants”

3. Construct a hypothesis: protein “Z” is a membrane receptor or interacts with a membrane receptor and participates in pathogen molecules recognition.

Scientific Method



4.Test hypothesis: This step involves designing multiple experiments to address each part of your question. This includes from basic experiments like checking for the presence of “z” in plant membranes or complex experiments like protein-protein interaction. Include positive and negative controls to avoid any false positive or false negatives in your experiment.

5. Data analysis: make use of statistical tools to make sense of the data you have obtained so far. Make relevant inferences, apply statistical tests and draw conclusions.

6. Results: report your conclusions in an understandable format which includes gel images, quantitative analysis, graphs etc. If “Z” shows a role in plant immunity, you can report you results in peer-reviewed journals, if not: either change the hypothesis, or check for errors in your analysis or try to report your findings as negative results.

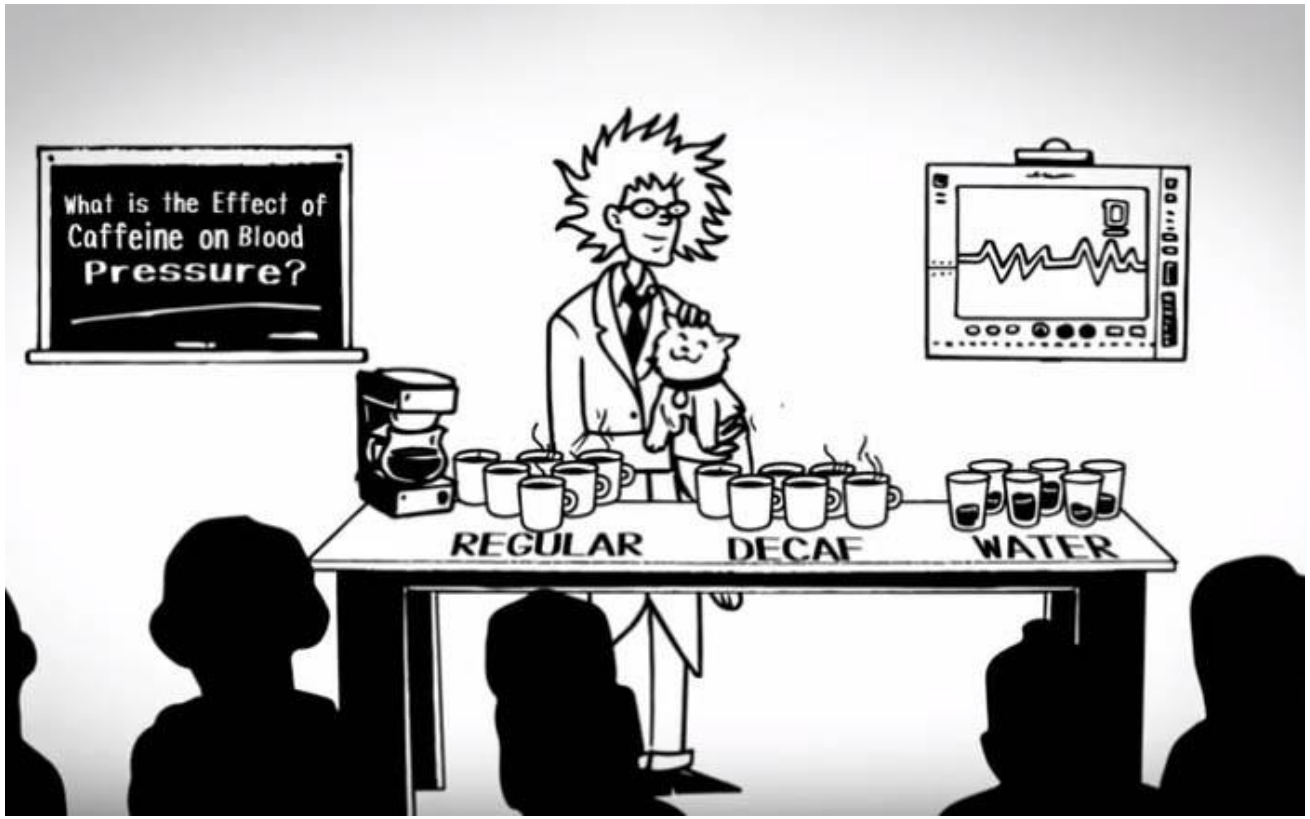
Scientific Method

4. Test hypothesis: This step involves **designing multiple experiments** to address each part of your question. This includes from basic experiments like checking for the presence of “z” in plant membranes or complex experiments like protein-protein interaction. Include positive and negative controls to avoid any false positive or false negatives in your experiment.

5. Data analysis: **make use of statistical tools to make sense of the data** you have obtained so far. Make relevant inferences, apply statistical tests and draw conclusions.

Experimental Design?

Experimental design



Experimental design includes both:

- Strategies for organizing data collection
- Because of variability: Data analysis procedures matched to those data collection strategies
- We wouldn't need a science of experimental design:
  - If all units (for e.g. organisms) were identical
  - If all units responded identically to treatments

- We need experimental design to control variability so that treatment effects can be identified

Experimental design

- The idea of controlling variability through design has a long history. One of the first exemples is the work in scruvy by Sir James Lind's (1747)

Without stating what method of allocation he used, Lind allocated two men to each of six different daily treatments for a period of fourteen days. The six treatments were:

1- 1.1 litres of cider;

2- twenty-five millilitres of elixir vitriol (dilute sulphuric acid);

3- 18 millilitres of vinegar three times throughout the day before meals;

4- half a pint of sea water;

5- two oranges and one lemon continued for six days only (when the supply was exhausted);

6- medicinal paste made up of garlic, mustard seed, dried radish root and gum myrrh.

Experimental design

"The most sudden and visible good effects were perceived from the use of oranges and lemons; one of those who had taken them being at the end of six days fit for duty … The other was the best recovered of any in his condition; and being now deemed pretty well, was appointed nurse to the rest of the sick."

Experimental design

- Later on.. Studies in crop variation (1921 – 1929) by statistician named Fisher was hired at Rothamsted agricultural station
- They had a lot of observational data on crop yields and hoped a statistician could analyze it to find effects of various treatments
- All he had to do was sort out the effects of confounding variables
- Fisher did regression analyses—lots of them—to study (and get rid of) the effects of confounders:
- soil fertility gradients
- drainage
- effects of rainfall
- effects of temperature and weather, etc.

Conclusion: The effects of confounders are typically larger than those of the systematic effects under study

Experimental design

- **Fisher invents**
  - Basic principles of experimental design
  - Control of variation by randomization
  - Analysis of variance

- **Studies in Crop variation IV (1927)**
  - Fisher invents analysis of covariance to combine statistical control and control by randomization

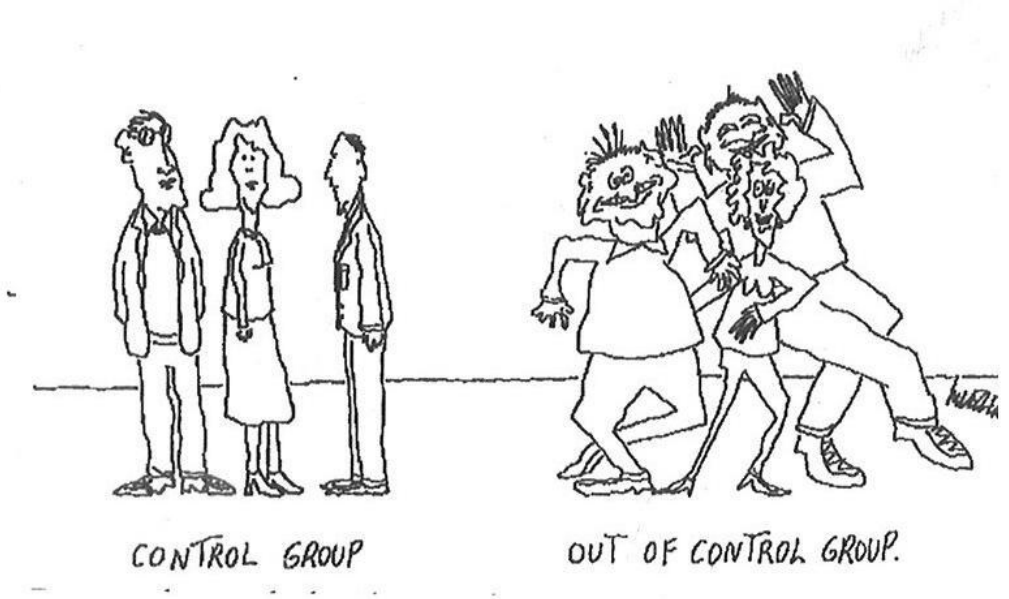
- **Studies in crop variation VI (1929)**
  - Fisher refines the theory of experimental design, introducing most other key concepts known today

Experimental design

- Experimental design controls background variability so that systematic effects of treatments can be observed

- Three basic principles
  - Control by matching (replication)
  - Control by randomization
  - Control by statistical adjustment

- Their importance is in that order

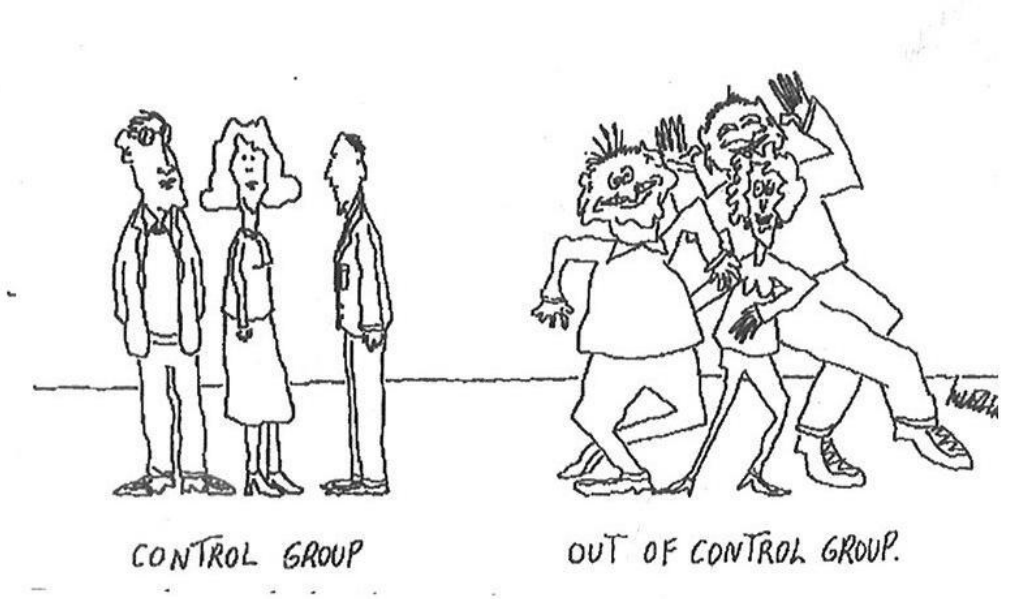


Experimental design

- Experimental design controls background variability so that systematic effects of treatments can be observed

- Three basic principles
  - Control by matching (replication)
  - Control by randomization
  - Control by statistical adjustment

- Their importance is in that order



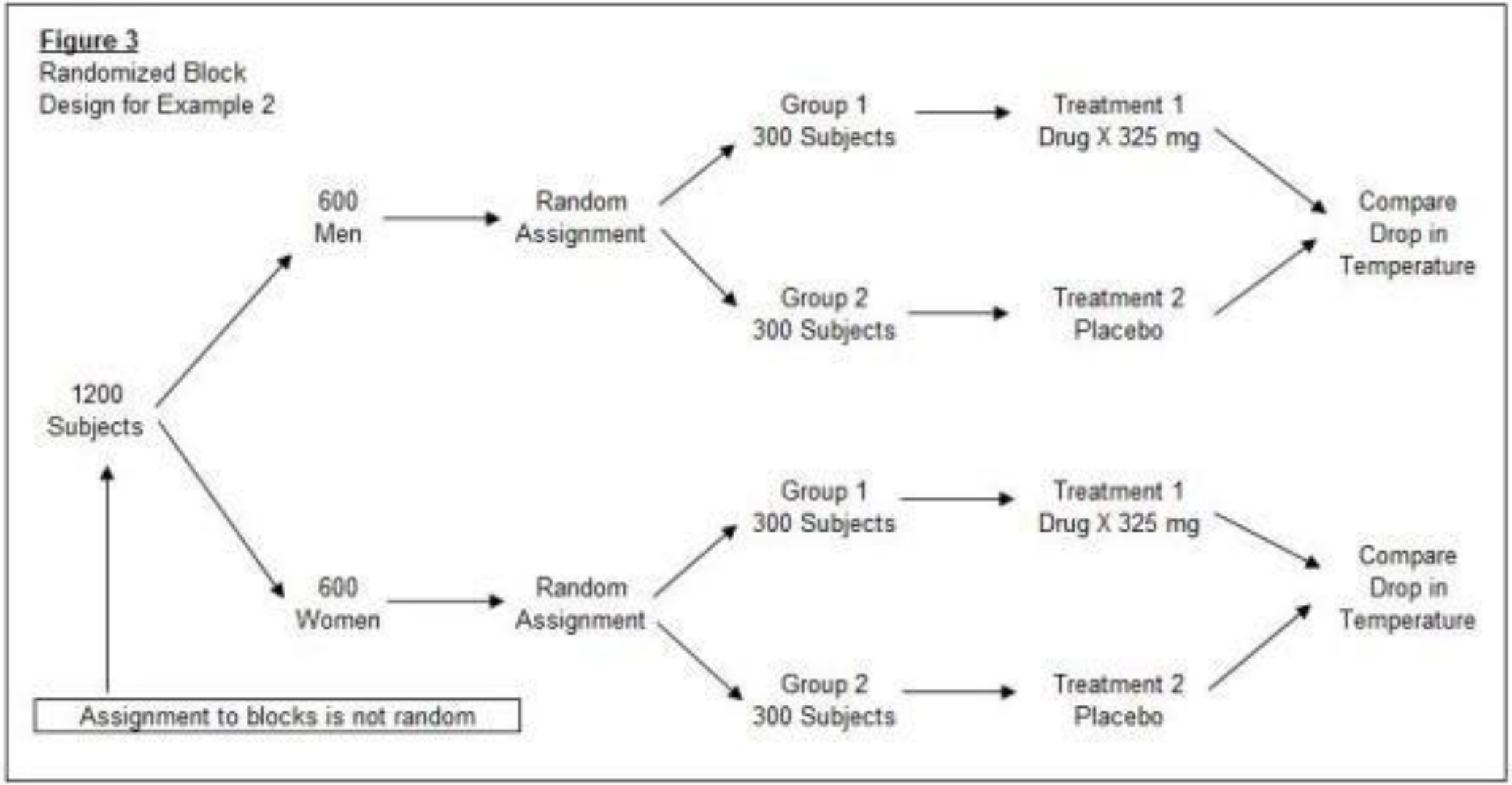
Experimental design

- Known sources of variation may be eliminated by matching (replication)

- However **matching is limited**
  - matching is only possible on observable characteristics
  - perfect matching is not always possible
  - matching inherently limits generalizability by removing (possibly desired) variation
- **Matching ensures that groups compared are alike on specific known and observable characteristics** (in principle, everything we have thought of)

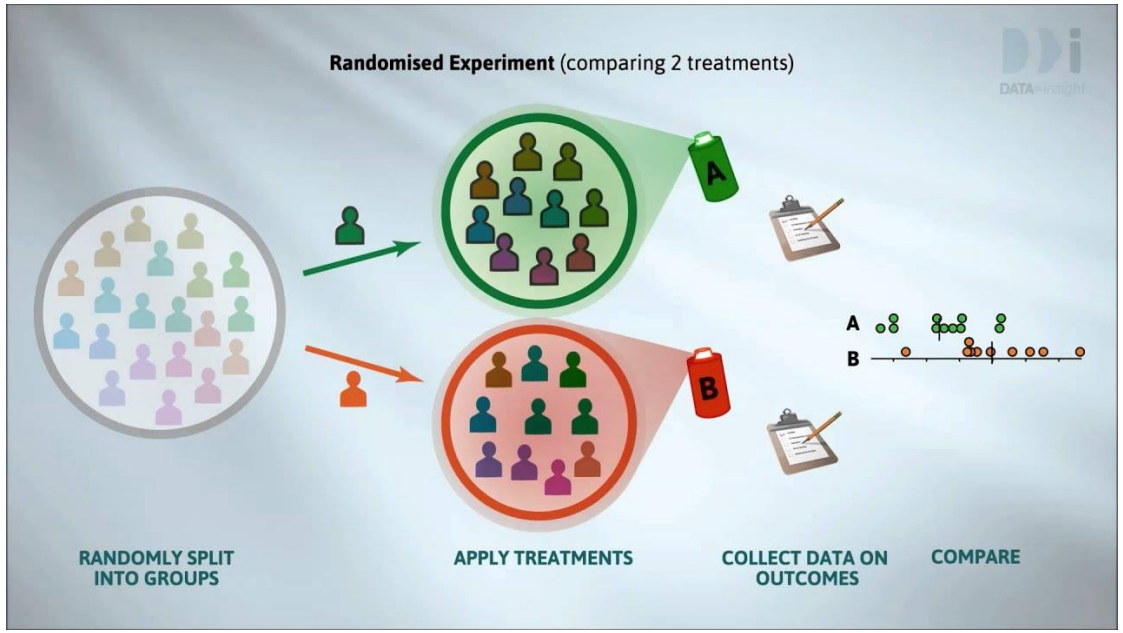
Wouldn't it be great if there were a method of making groups alike on not only everything we have thought of, but everything we didn't think of too?

Experimental design



Experimental design

- Randomization controls for the effects all (observable or non-observable, known or unknown) characteristics
- Randomization makes groups equivalent (on average) on all variables (known and unknown, observable or not)
- Randomization also gives us a way to assess whether differences after treatment are larger than would be expected due to chance.
- Random assignment is not assignment with no particular rule. It is a purposeful process



Experimental design

Why are experiments conducted?

Experimental design

Why are experiments conducted?

Experiments are used to study causal relationships - you later independent variables and measure the alteration effect on the dependent variables

Experimental design

Why are experiments conducted?

Experiments are used to study causal relationships - you later independent variables and measure the alteration effect on the dependent variables

How about experimental design?

Experimental design

Why are experiments conducted?

Experiments are used to study causal relationships - you later independent variables and measure the alteration effect on the dependent variables

How about experimental design?

Set of procedures to systematically test a hypothesis. A good experimental design requires a strong understanding of the system you are studying.

Experimental design

Why are experiments conducted?

Experiments are used to study causal relationships - you later independent variables and measure the alteration effect on the dependent variables

How about experimental design?

Set of procedures to systematically test a hypothesis. A good experimental design requires a strong understanding of the system you are studying.

Where to start?

1- define your variables

2- define your hypothesis

3-design experimental treatments to manipulate the independent variables

4-Assign subjects to groups, either between-subjects or within-subjects (matching and randomization)

5- Think of how to measure the effects in the dependent variables

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis?

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis?

Define your variables!

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis? How to define it?

Independent variable - minutes of phone use before sleep

Dependent variable - hours of sleep per night

Research question	Independent variable	Dependent variable
Phone use and sleep	Minutes of phone use before sleep	Hours of sleep per night

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis? How to define it?

Independent variable - minutes of phone use before sleep

Dependent variable - hours of sleep per night

What are the external factors influencing your experiment?

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis? How to define it?

Independent variable - minutes of phone use before sleep

Dependent variable - hours of sleep per night

What are the external factors influencing your experiment?

Individual sleep patterns - how to overcome this?

Experimental design

Examples:

Phone use and sleep

Question?

How does phone use before bedtime affects sleep patterns.

Can you further specify your question?

How the number of minutes a person used their phone before sleep affects the number of sleep hours

What is your experimental hypothesis? How to define it?

Independent variable - minutes of phone use before sleep

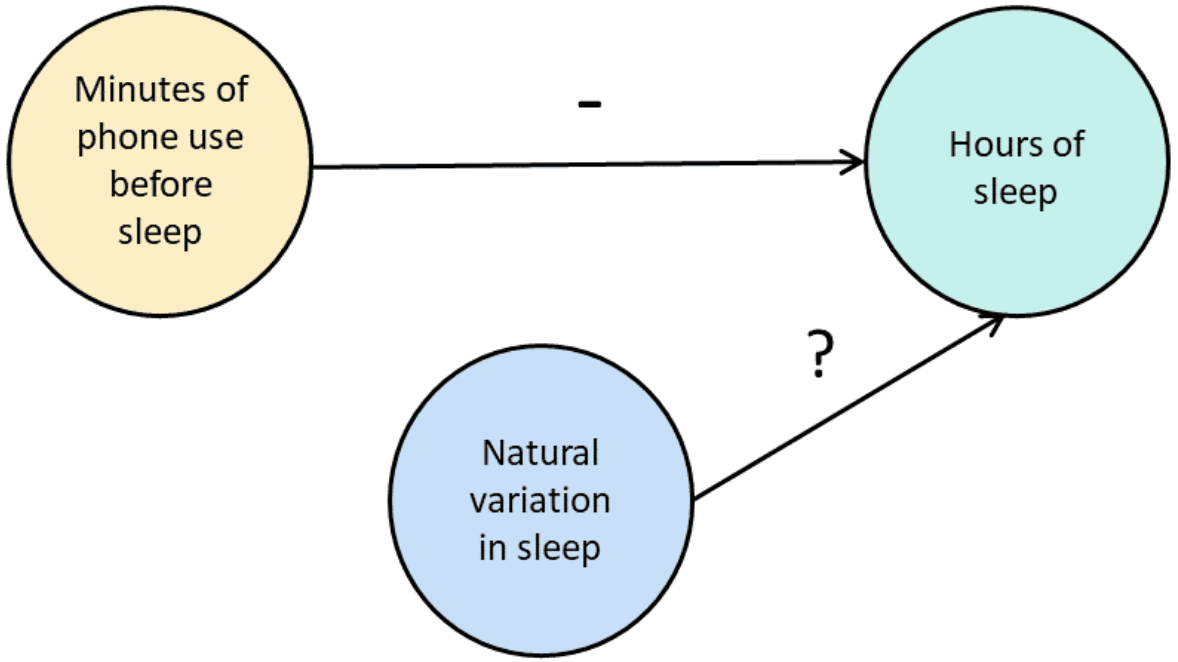
Dependent variable - hours of sleep per night

What are the external factors influencing your experiment?

Individual sleep patterns - how to overcome this?

-measure the difference between sleep with phone use and sleep without phone use in the groups (rather than the average amount of sleep per treatment group)

Experimental design



Experimental design

Write your hypothesis:

Experimental design

Write your hypothesis:

H0: phone use before sleep does not correlate with the amount of sleep a person gets

H1: Increasing phone use before sleep leads to a decrease in sleep.

Define a controlled experimental approach

Experimental design

Write your hypothesis:

H0: phone use before sleep does not correlate with the amount of sleep a person gets

H1: Increasing phone use before sleep leads to a decrease in sleep.

Define a controlled experimental approach

Design your experimental treatments and groups

Use the phone as a:

- categorical variable: either as binary (yes/no) or as levels of a factor (no phone use, low phone use, high phone use).

- continuous variable (minutes of phone use measured every night).

Experimental design

Write your hypothesis:

H0: phone use before sleep does not correlate with the amount of sleep a person gets

H1: Increasing phone use before sleep leads to a decrease in sleep.

Define a controlled experimental approach

Design your experimental treatments and groups

Use the phone as a:

- categorical variable: either as binary (yes/no) or as levels of a factor (no phone use, low phone use, high phone use).

- continuous variable (minutes of phone use measured every night).

Define study size

-How many individuals? Control group? How to define a “control”

-Statistical power?

Experimental design

Define study size

-How many individuals? Control group? How to define a “control”

-Statistical power?

- *Randomized design or randomized block design?*
- *Between subjects design or within subjects design?*

Experimental design

Define study size

-How many individuals? Control group? How to define a “control”

-Statistical power?

- *Randomized design or randomized block design?*
- *Between subjects design or within subjects design?*

	Completely randomized design	**Randomized block design**
Phone use and sleep	Subjects are all randomly assigned a level of phone use using a random number generator.	Subjects are first grouped by age, and then phone use treatments are randomly assigned within these groups.

Experimental design

Measure your dependent variable

- *How to do this? How to operationalize it?*
- *How to get robust data?*

	Between-subjects (independent measures) design	Within-subjects (repeated measures) design
Phone use and sleep	Subjects are randomly assigned a level of phone use (none, low, or high) and follow that level of phone use throughout the experiment.	Subjects are assigned consecutively to zero, low, and high levels of phone use throughout the experiment, and the order in which they follow these treatments is randomized.

Experimental design

Measure your dependent variable

- *How to do this? How to operationalize it?*
- *How to get reliable and valid measurments that minimize research bias or error?*
- *How to get robust data?*

suggestions:

ask participants to record the time they go to sleep and get up,

Ask participants to wear a sleep tracker

Challenge

To define an experimental design

"My team has created a very innovative solution, but we're still looking for a problem to go with it."

GPS LAB
Grapevine Pathogen Systems Lab

Speaking the language of lipids in grapevine-*Plasmopara viticola* interaction

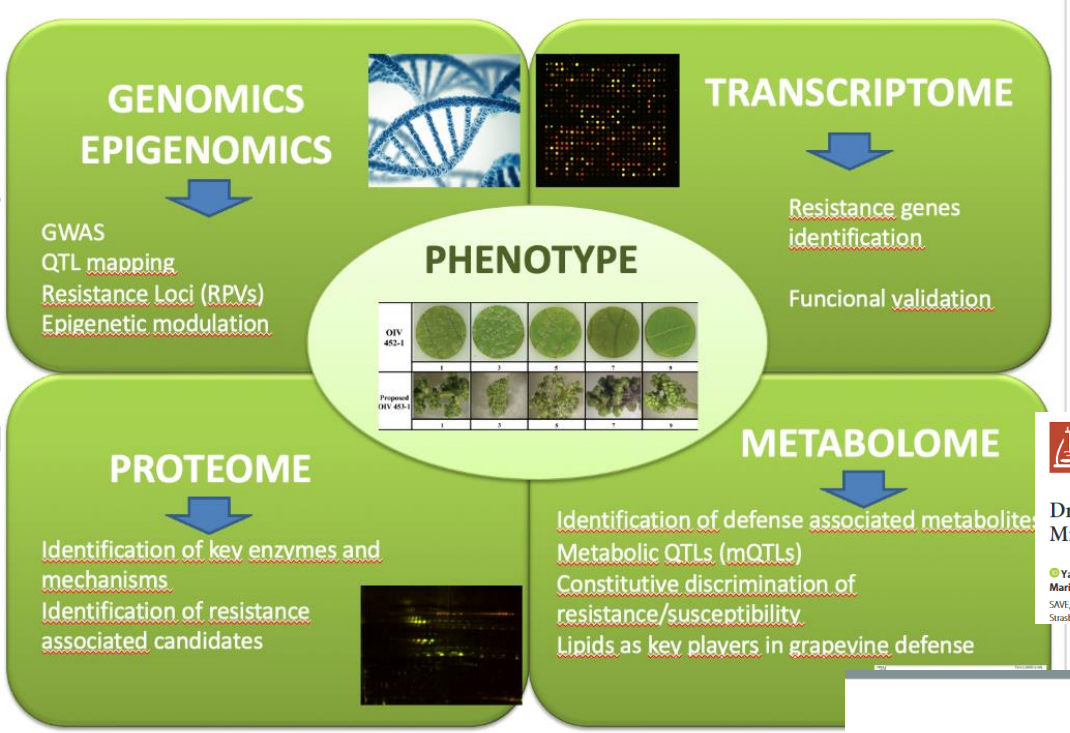
BioISI
[integrating sciences]

Ciências ULisboa
Faculdade de Ciências da Universidade de Lisboa

Grapevine downy mildew research

Fifteen years of molecular approaches for disease and resistance mechanisms disclosure

Vitis vinifera
2007 -Grapevine first genome dr
2014 - Genome re-annotation



Plasmopara viticola
Obligate biotroph
2012 - Sequencing projects – INRA
2016 -First genome draft
2019 - Genome available


Vol 449|27 September 2007|doi:10.1038/nature06148 nature

LETTERS

The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla

The French–Italian Public Consortium for Grapevine Genome Characterization*

The analysis of the first plant genomes provided unexpected evidence for genome duplication events in species that had previously been considered as true diploids on the basis of their genetics[1,2]. ...

All grapevine varieties are highly heterozygous; preliminary data showed that there was as much as 13% sequence divergence between alleles, which would hinder reliable contig assembly when a whole-genome shotgun strategy was used for sequencing. Our consortium ...


OPEN ACCESS Freely available online PLoS one

A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety

Riccardo Velasco[1*], Andrey Zharkikh[2*], Michela Troggio[1], Dustin A. Cartwright[1,2], Alessandro Cestaro[1], Dmitry Pruss[2], Massimo Pindo[1], Lisa M. FitzGerald[2], Silvia Vezzulli[1], Julia Reid[2], Giulia Malacarne[1], Diana Iliev[2], Giuseppina Coppola[1], Bryan Wardell[2], Diego Micheletti[1], Teresita Macalma[2], Marco Facci[1], Jeff T. Mitchell[2], Michele Perazzolli[1], Glenn Eldredge[2], Pamela Gatto[1], Rozan Oyzerski[2], Marco Moretto[1], Natalia Gutin[2], Marco Stefanini[1], Yang Chen[2], Cinzia Segala[1], Christine Davenport[2], Lorenzo Dematté[1], Amy Mraz[2], Juri Battilana[1], Keith Stormo[2], Fabrizio Costa[1], Quanzhou Tao[2], Azeddine Si-Ammour[1], Tim Harkins[2], Angie Lackey[2], Clotilde Perbost[2], Bruce Taillon[2], Alessandra Stella[2], Victor Solovyev[2], Jeffrey A. Fawcett[2], Lieven Sterck[2], Klaas Vandepoele[2], Stella M. Grando[1], Stefano Toppo[2], Claudio Moser[1], Jerry Lanchbury[2], Robert Bogden[2], Mark Skolnick[2], Vittorio Sgaramella[2], Satish K. Bhatnagar[2], Paolo Fontana[1], Alexander Gutin[2], Yves Van de Peer[2], Francesco Salamini[1], Roberto Viola[1]

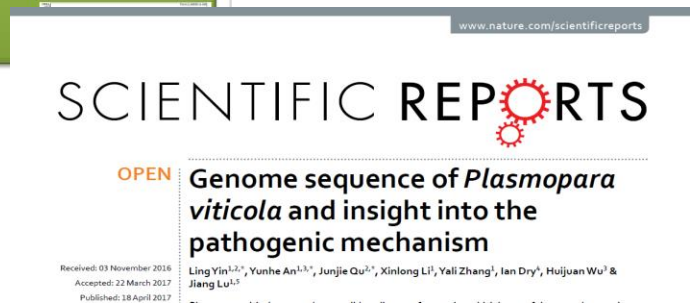
1 IASMA Research Center, San Michele all'Adige, Trento, Italy, 2 Myriad Genetics Inc, Salt Lake City, Utah, United States of America, 3 454 Life Sciences Corporation, Branford, Connecticut, United States of America, 4 Roche Diagnostics Corporation, Roche Applied Science, Indianapolis, Indiana, United States of America, 5 Amplicon Express Inc., Pullman, Washington, United States of America, 6 Technology Park Lodi, Lodi, Italy, 7 Department of Plant Systems Biology, VIB, Gent University, Gent, Belgium, 8 Department of Biological Chemistry, Padova University, Padova, Italy, 9 Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, United Kingdom


AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of *Plasmopara viticola*, the Grapevine Downy Mildew Pathogen

Yann Dussert,[a] Jérôme Gouzy,[b] Sylvie Richart-Cervera,[a] Isabelle D. Mazet,[a] Laurent Delière,[a] Carole Couture,[a] Ludovic Legrand,[b] Marie-Christine Piron[c] Pere Mestre,[c] François Delmotte[a]

SAVE, Bordeaux Sciences Agro, INRA, Villenave d'Ornon, France[a]; LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France[b]; SVQV, INRA, Université de Strasbourg, Colmar, France[c]


www.nature.com/scientificreports

SCIENTIFIC REPORTS

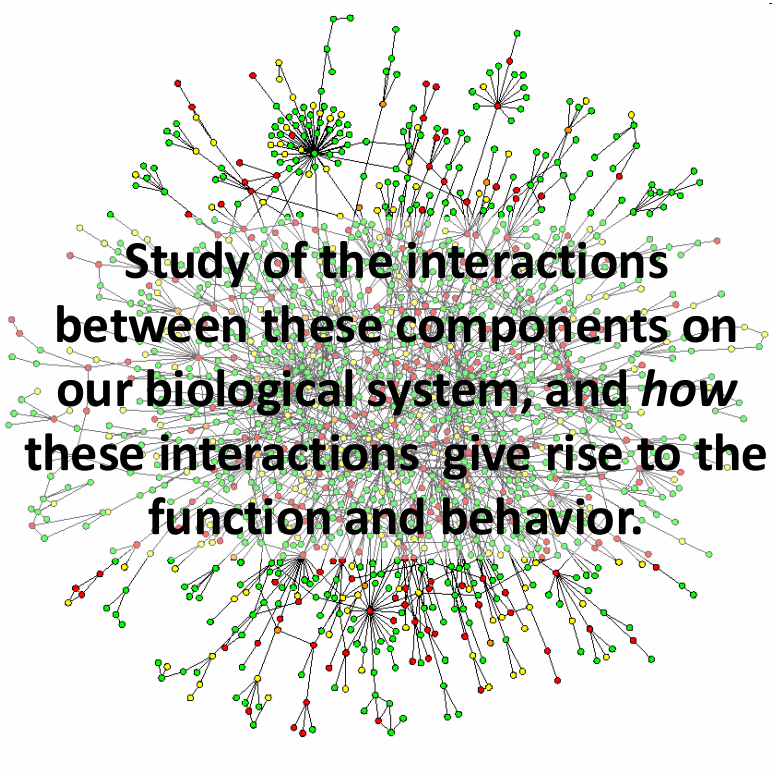
OPEN

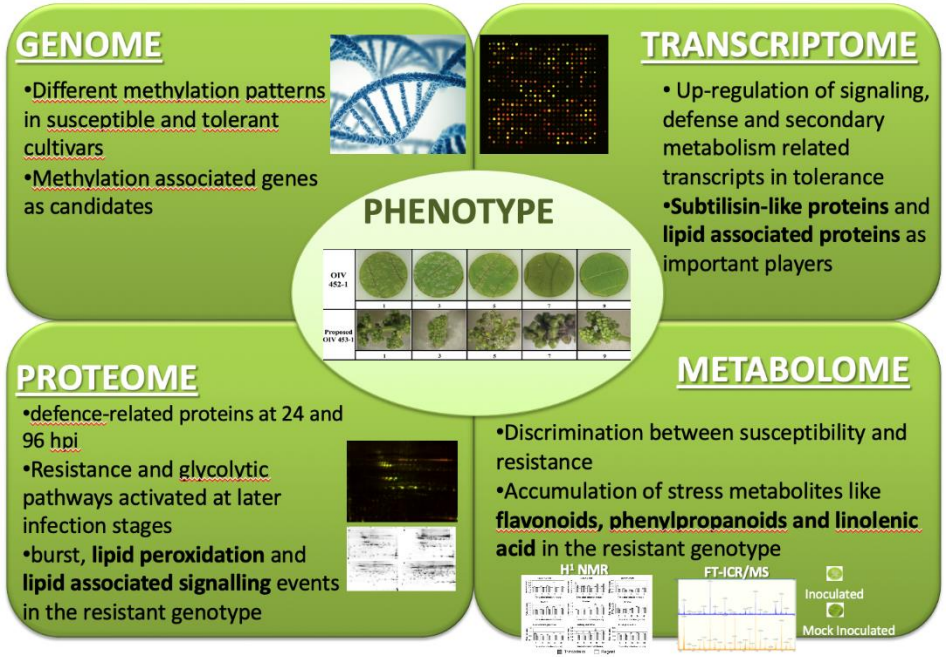
Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism

Received: 03 November 2016
Accepted: 22 March 2017
Published: 18 April 2017

Ling Yin[1,2,*], Yunhe An[1,3,*], Junjie Qu[1,*], Xinlong Li[1], Yali Zhang[1], Ian Dry[4], Huijuan Wu[1] & Jiang Lu[1,5]

Systems biology – something about networks

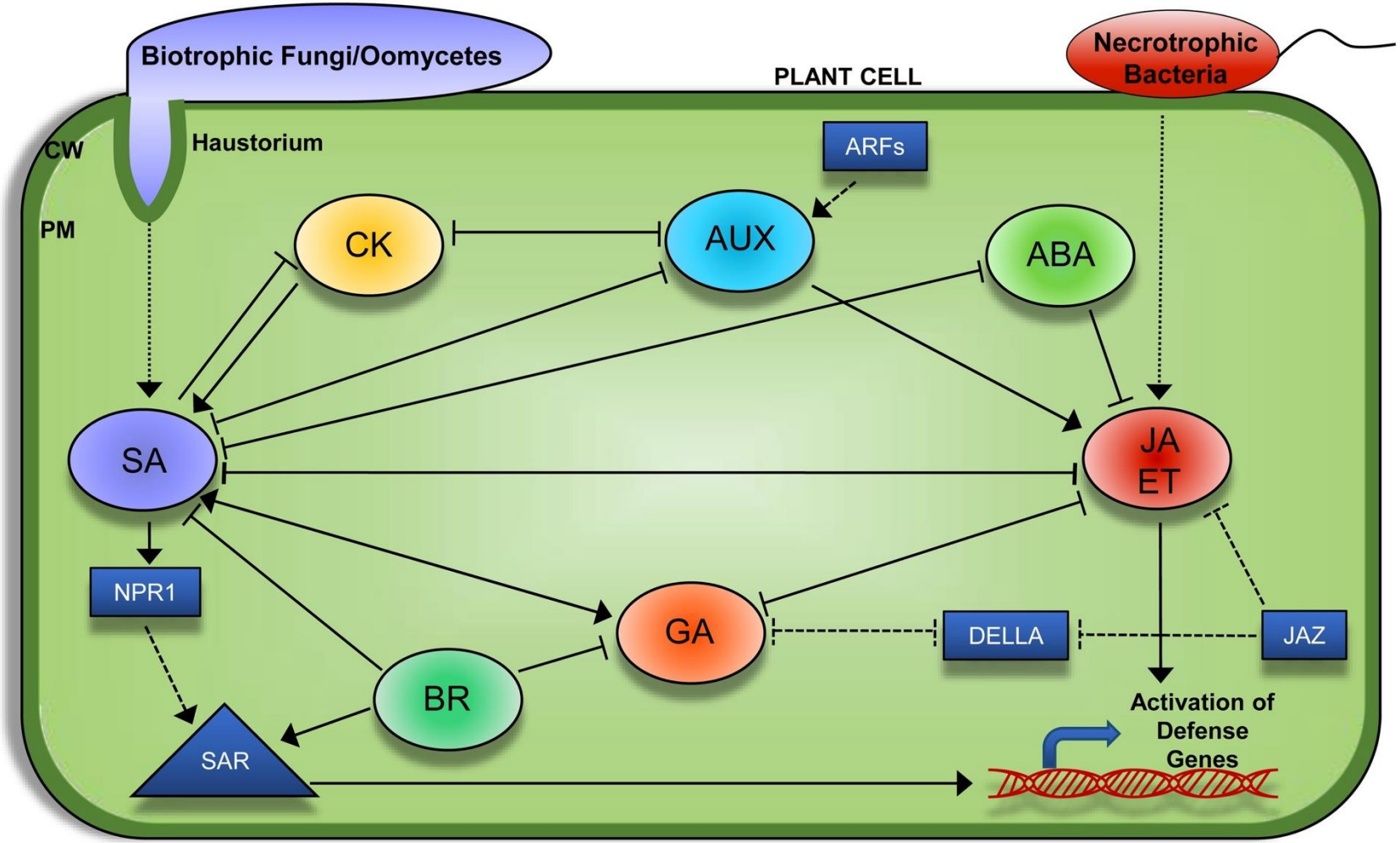




Jasmonic acid signaling? Lipids? Fatty acids?

Hormone signaling in plant-pathogen interactions

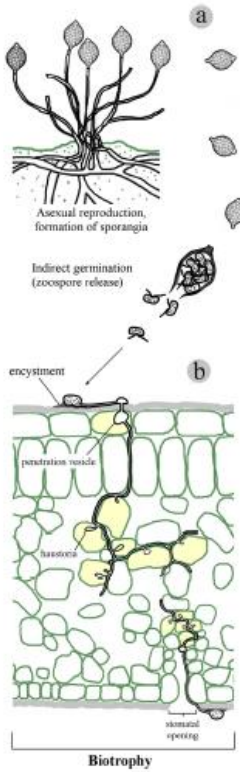
What was the basic knowledge on hormone signaling in plant-pathogen interactions?

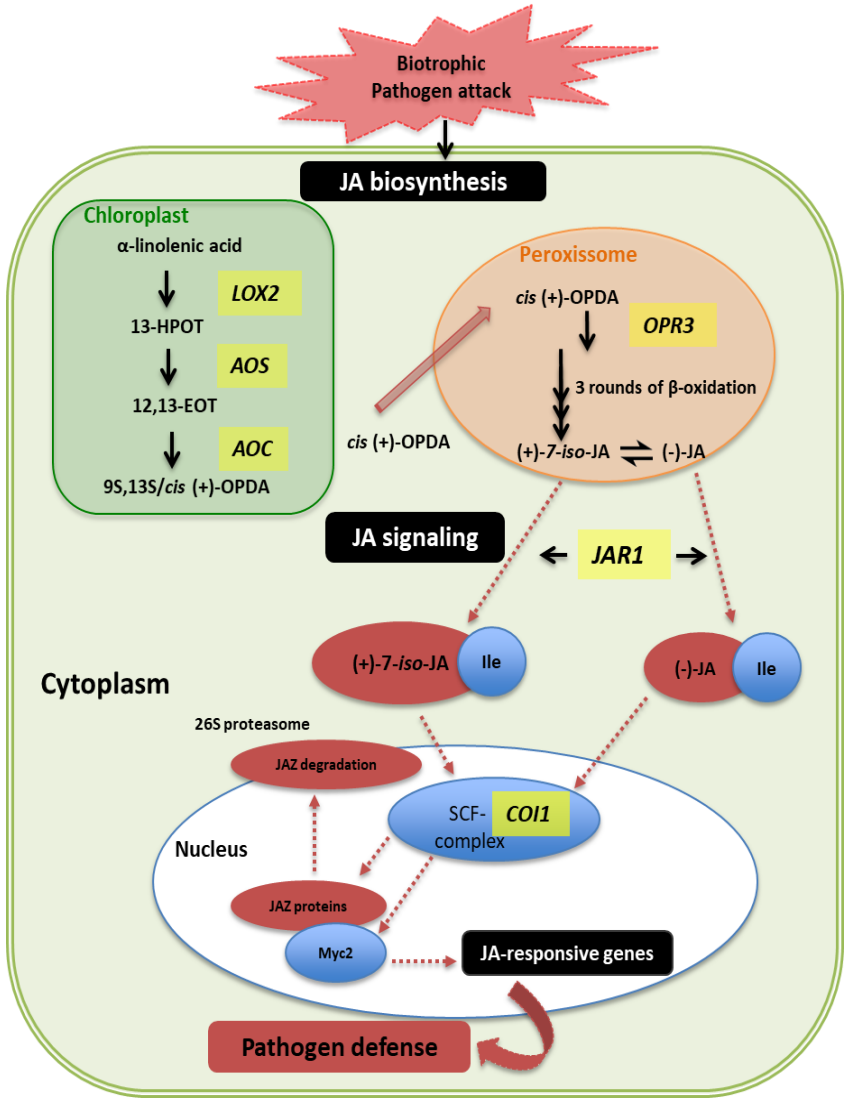


https://doi.org/10.1016/j.semcdb.2016.06.005

Hormone signaling in grapevine-P. viticola interaction

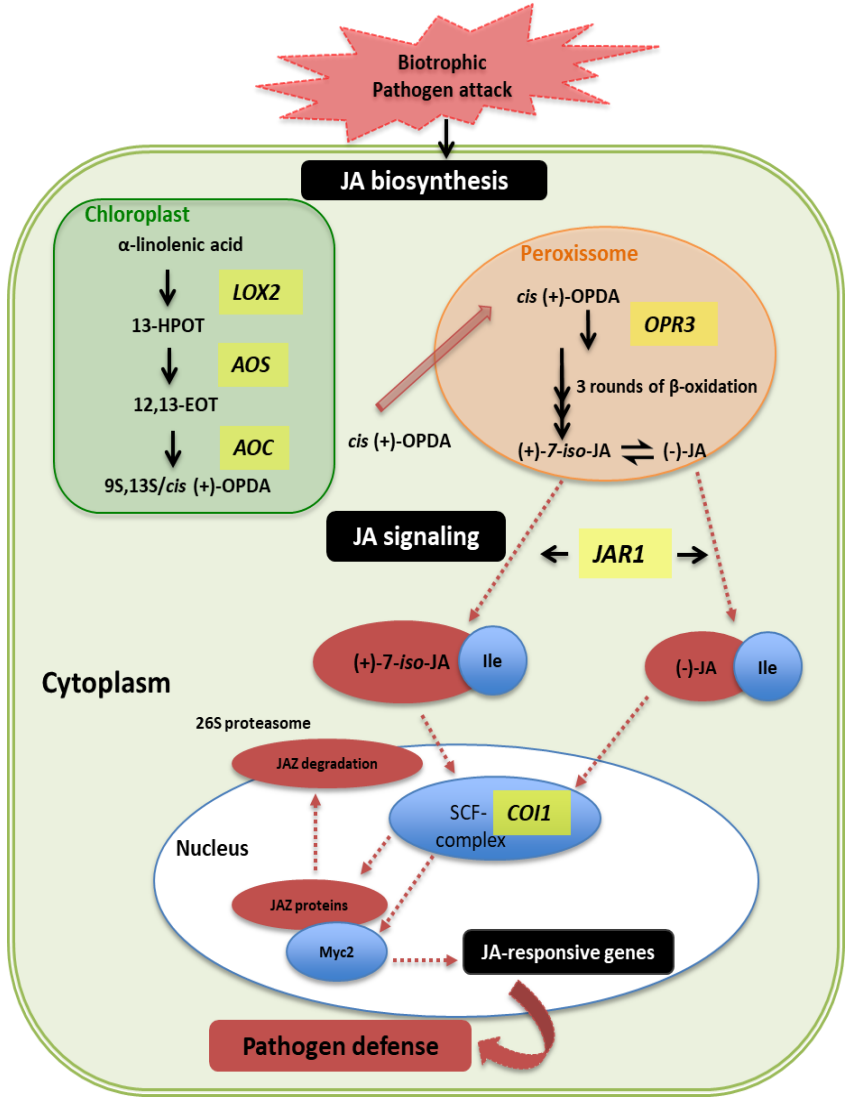
Jasmonic acid signaling activated in the first hours of infection with a biotroph?

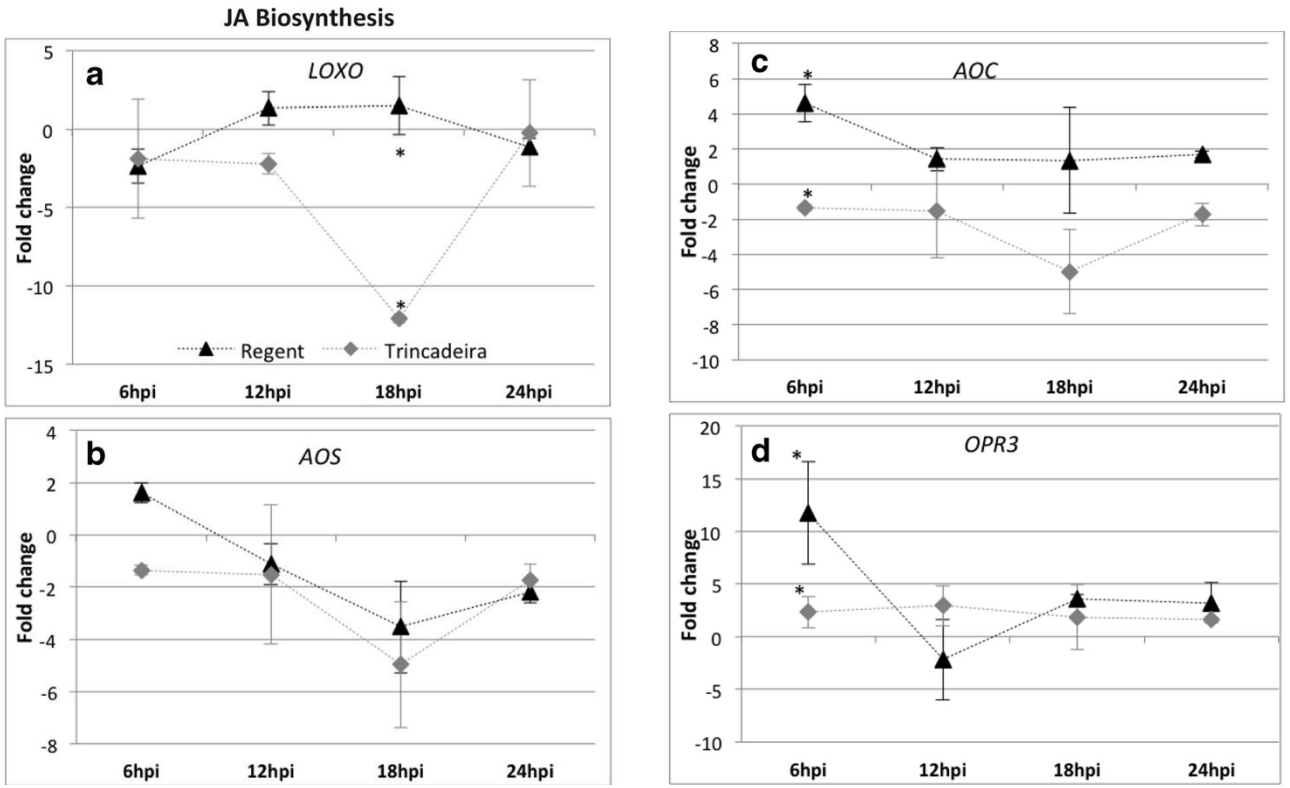




Hormone signaling in grapevine-P. viticola interaction

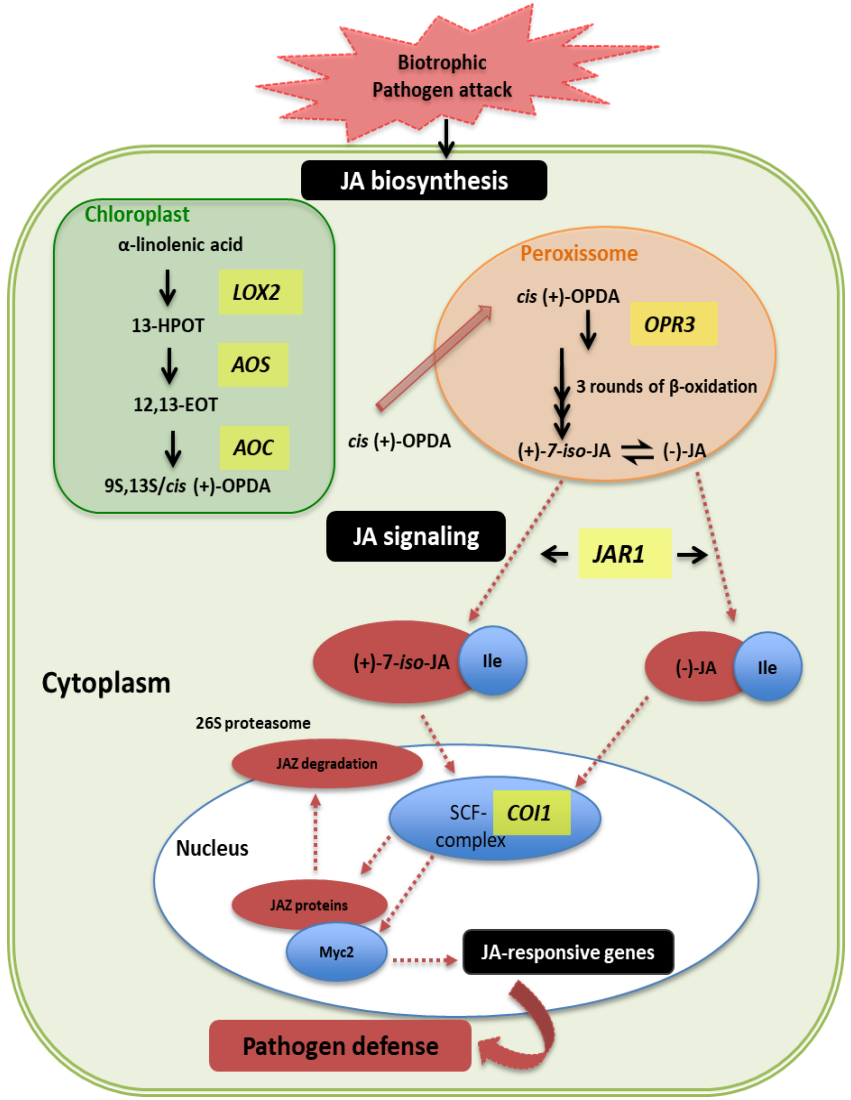
Jasmonic acid signaling activated in the first hours of infection with a biotroph?

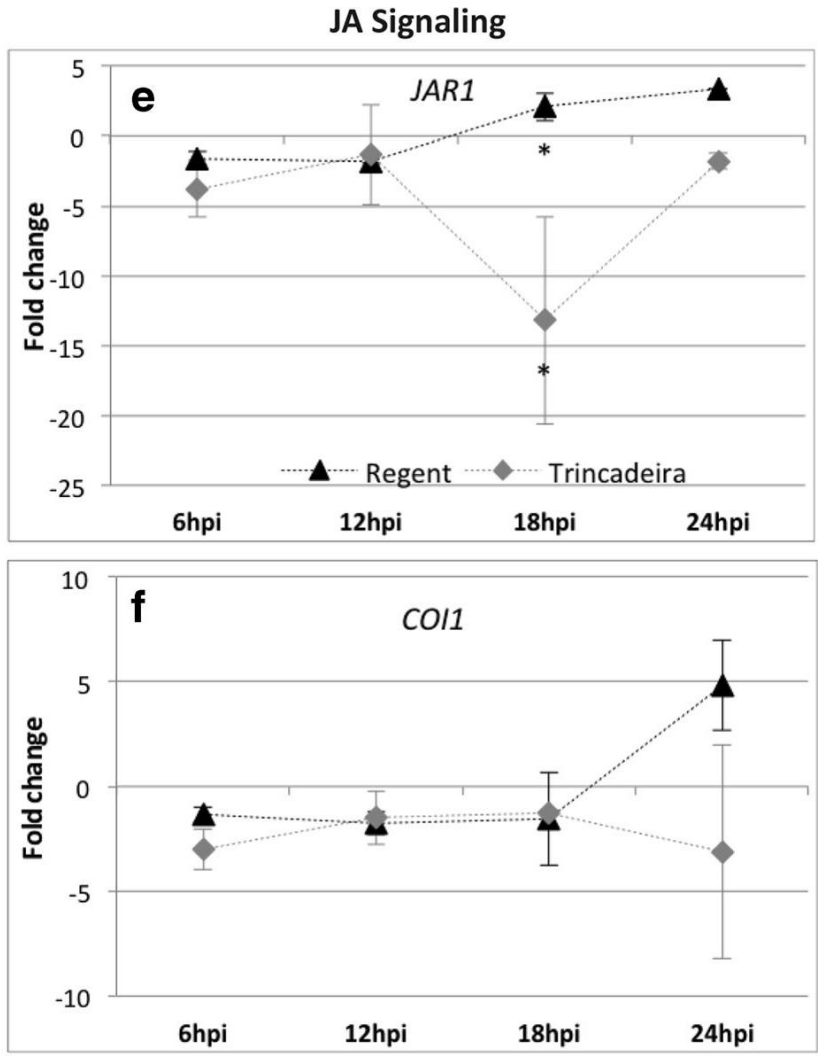




Hormone signaling in grapevine-P. viticola interaction

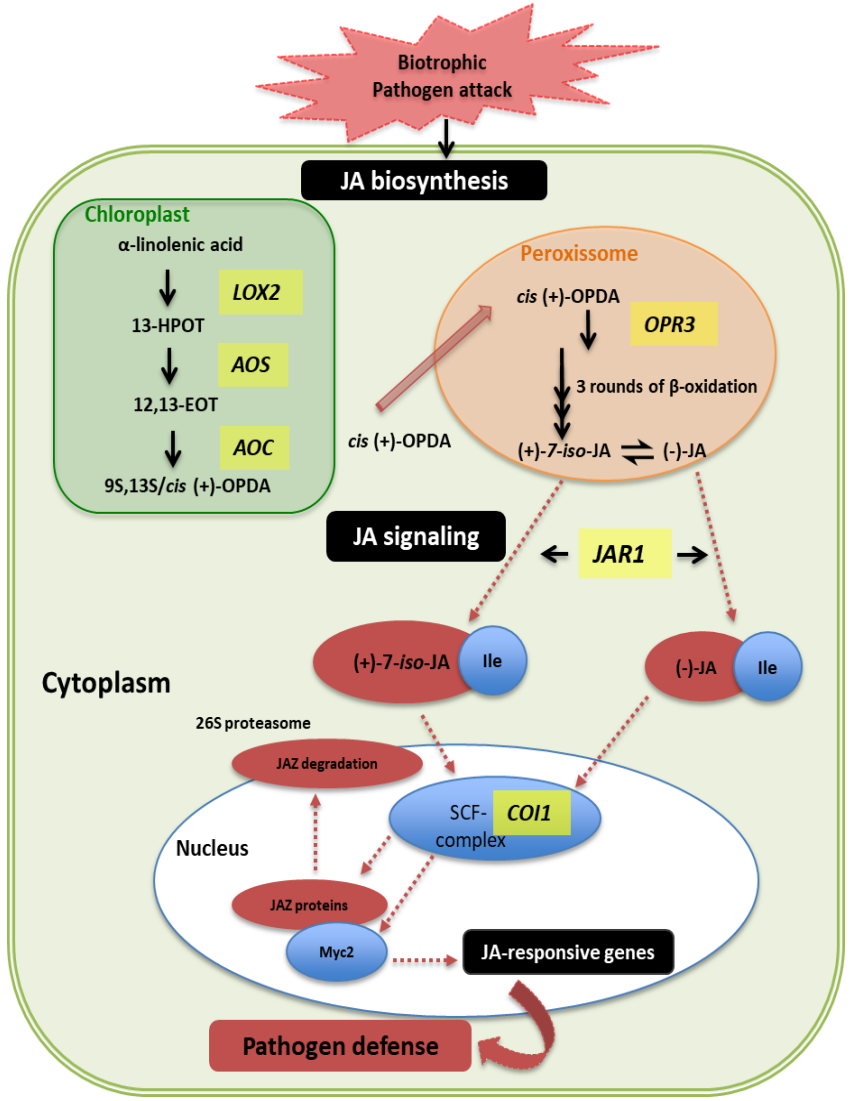
Jasmonic acid signaling activated in the first hours of infection with a biotroph?

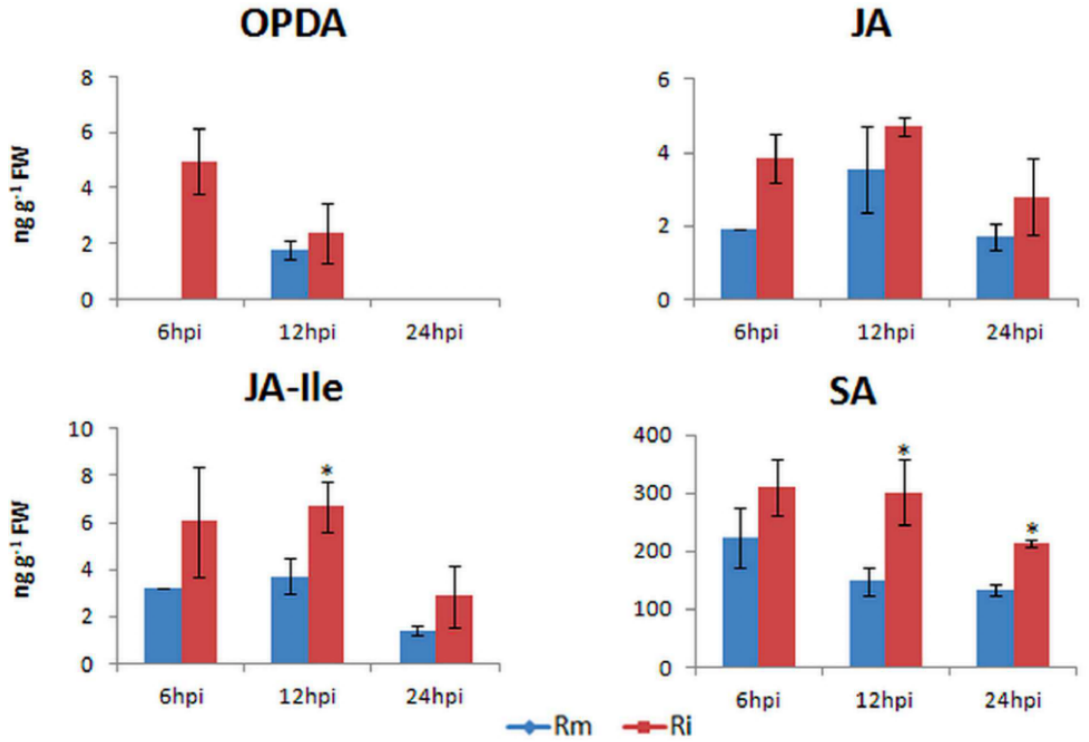




Hormone signaling in grapevine-P. viticola interaction

Jasmonic acid signaling activated in the first hours of infection with a biotroph?



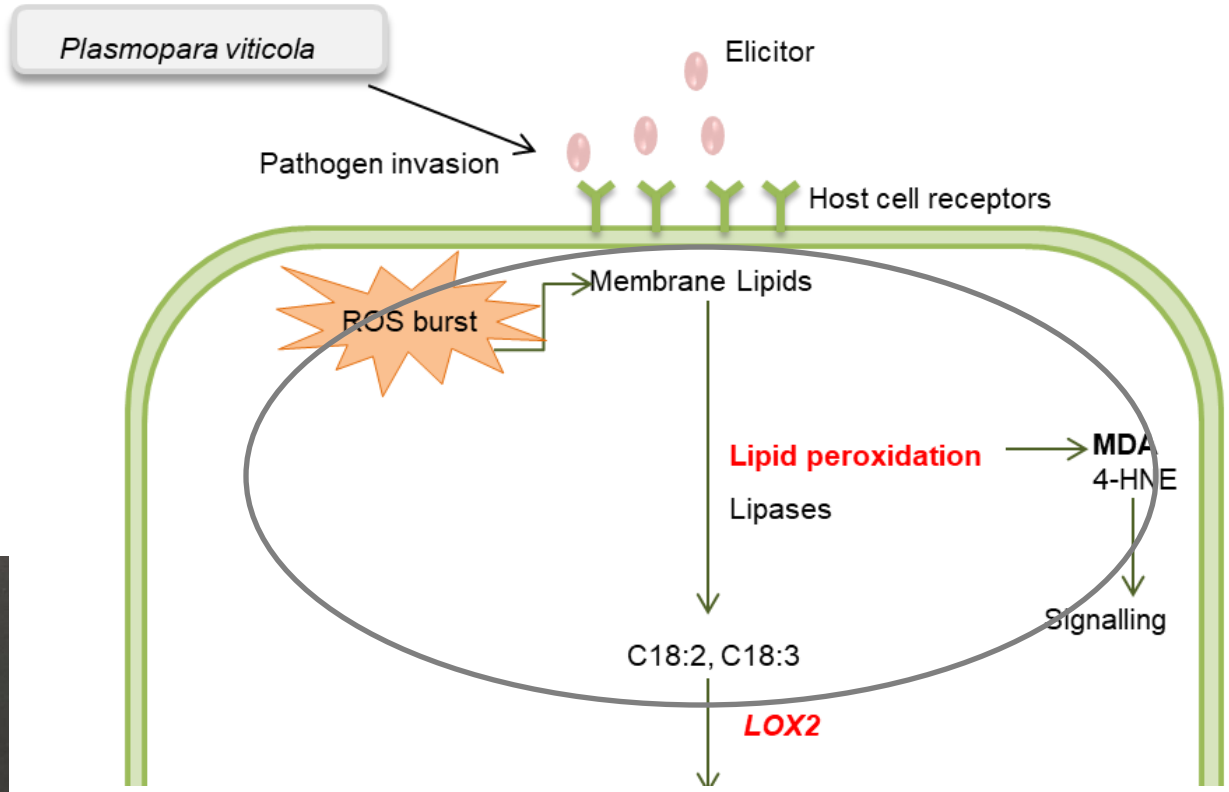


Figueiredo et al. (2015) *European Journal of Plant Pathology*
Guerreiro et al. (2016) *Frontiers in Plant Science*
Laureano et al. (2018) *Scientific Reports*
Cavaco et al. (2019) ISHS

Which events lead to JA precursor accumulation?

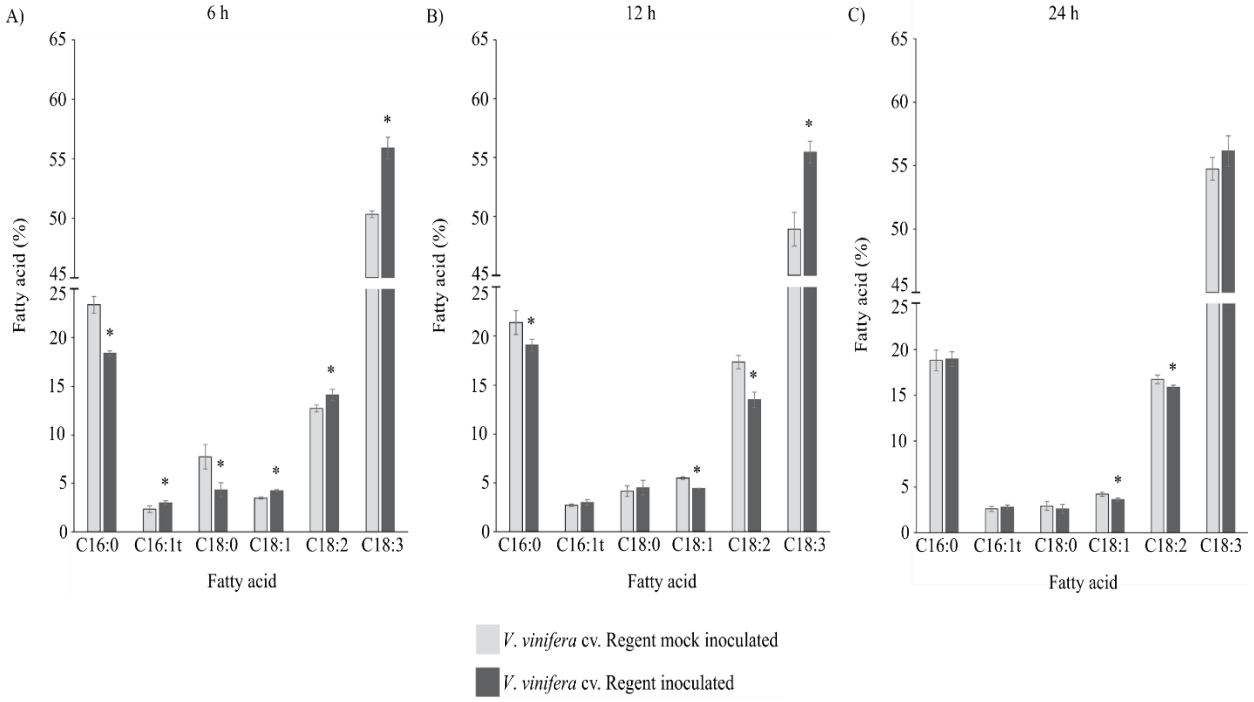
Tolerant cultivar:

↑ **JA biosynthesis**
↑ Accumulation of the precursor **C18:3**



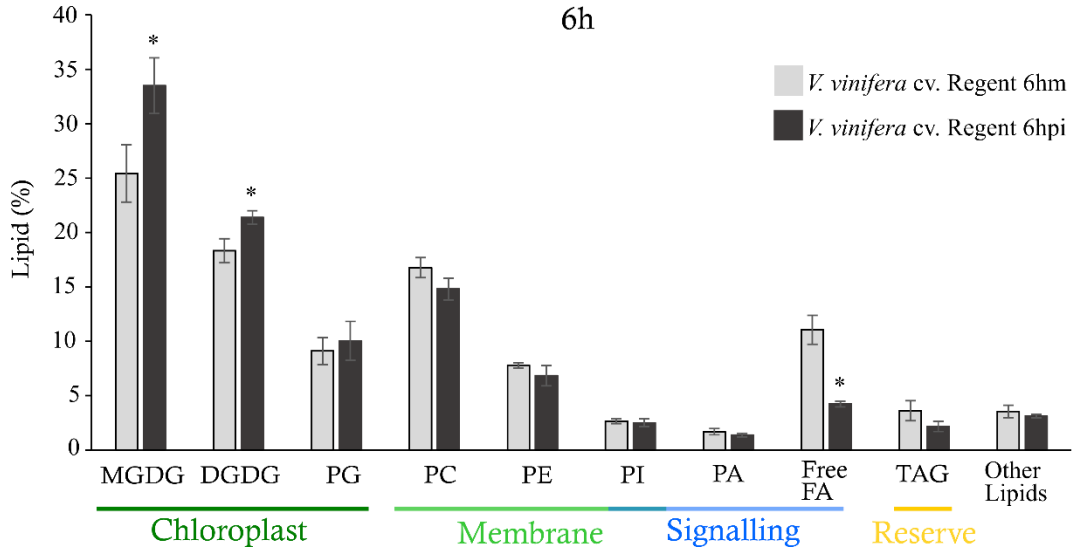
Why and how tolerant cultivars present a higher content of free C18:3 in the chloroplast?

Which events lead to JA precursor accumulation?

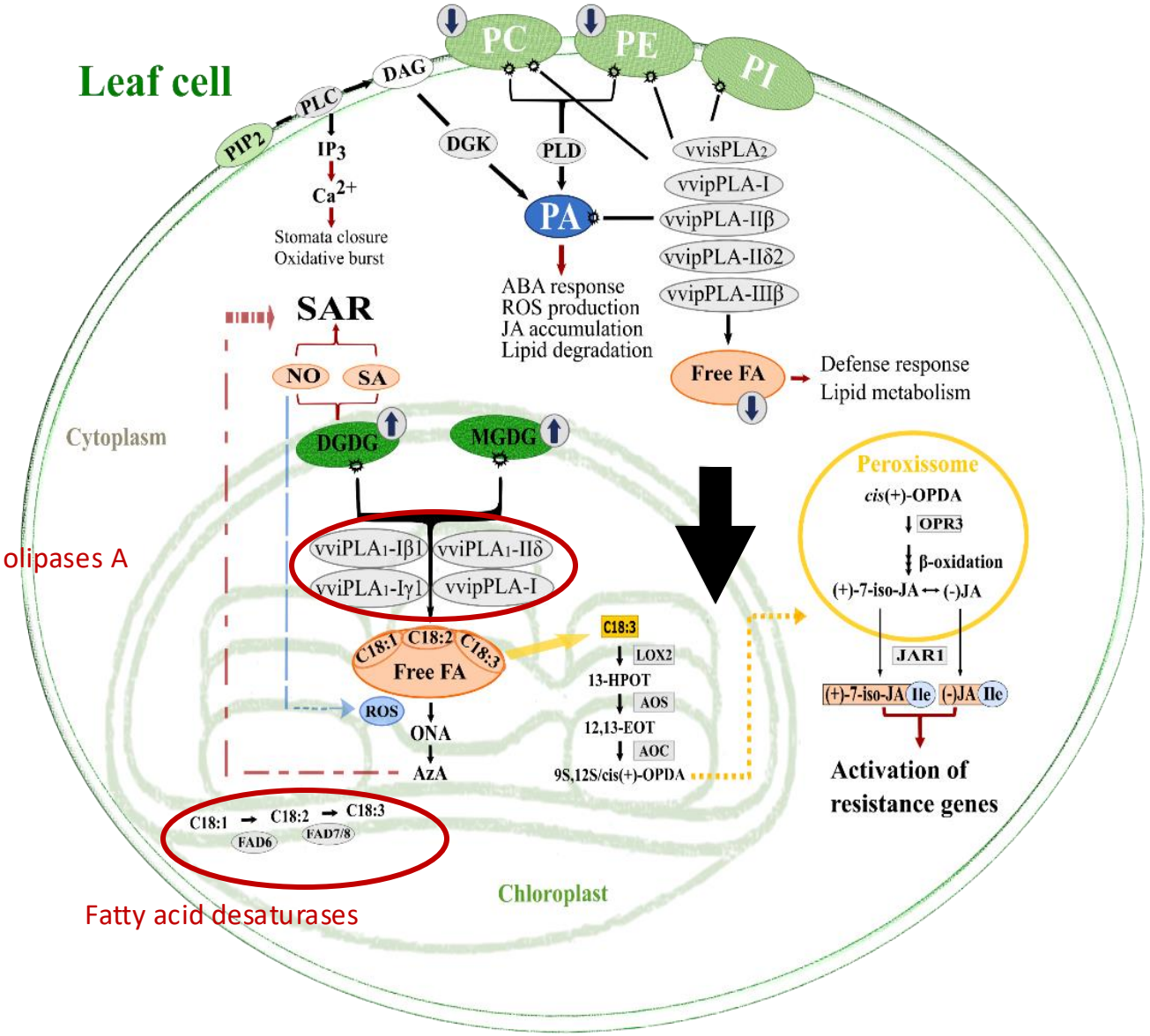


FA modulation after *P. viticola* inoculation **only occurs in the tolerant cultivar**

Increase of C18 desaturation;
C18:3 accumulation;
Increase of PUFAs.



P. viticola inoculation triggers FA and lipid content alteration



Players

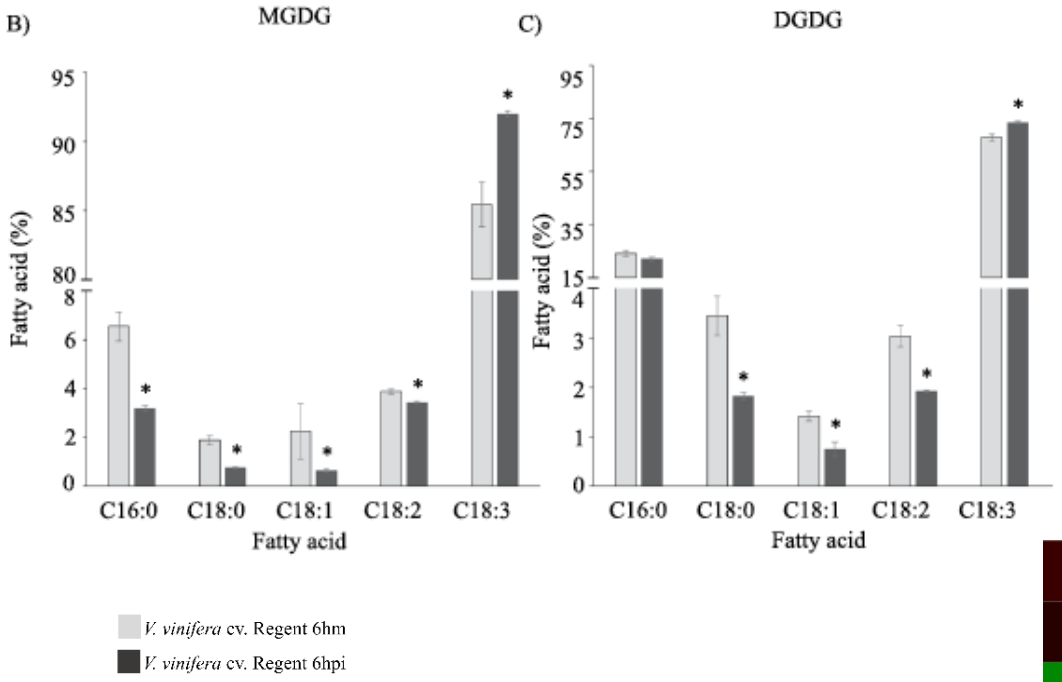
Phospholipases

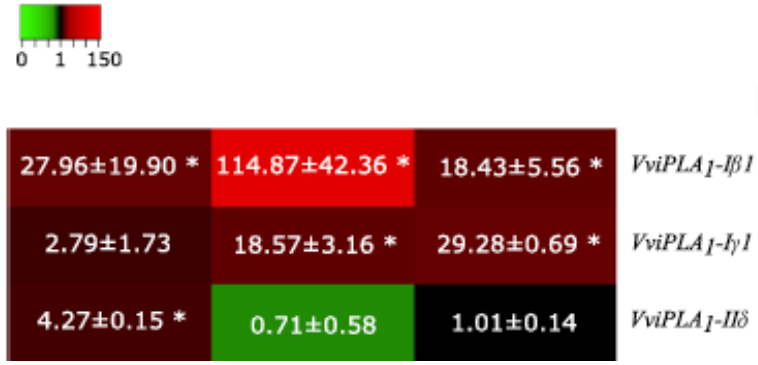
Catalyze the hydrolysis membrane lipids into lysophospholipids and fatty acids

Fatty acid desaturases

Introduce double bonds in fatty acids present in membrane lipids leading to the synthesis of unsaturated FA

P. viticola inoculation triggers FA and lipid content alteration

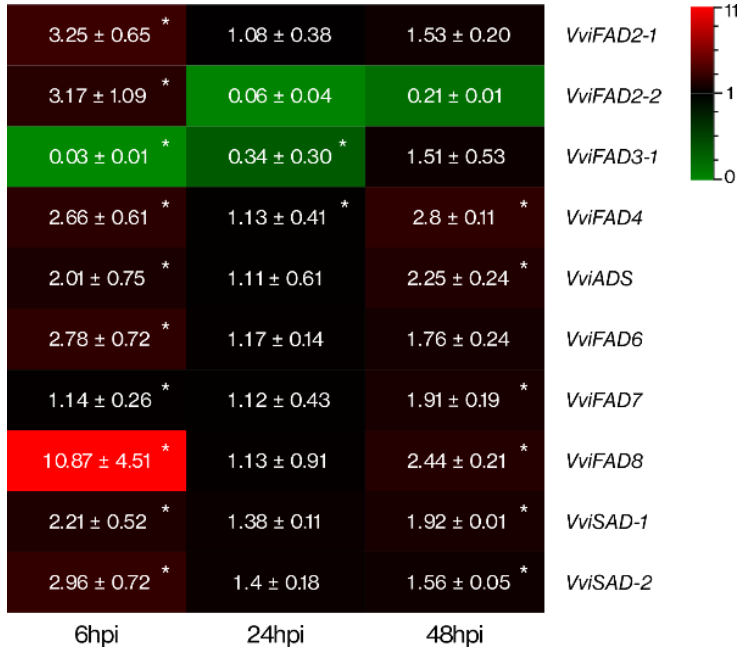




Activation of chloroplastidial Phospholipases A

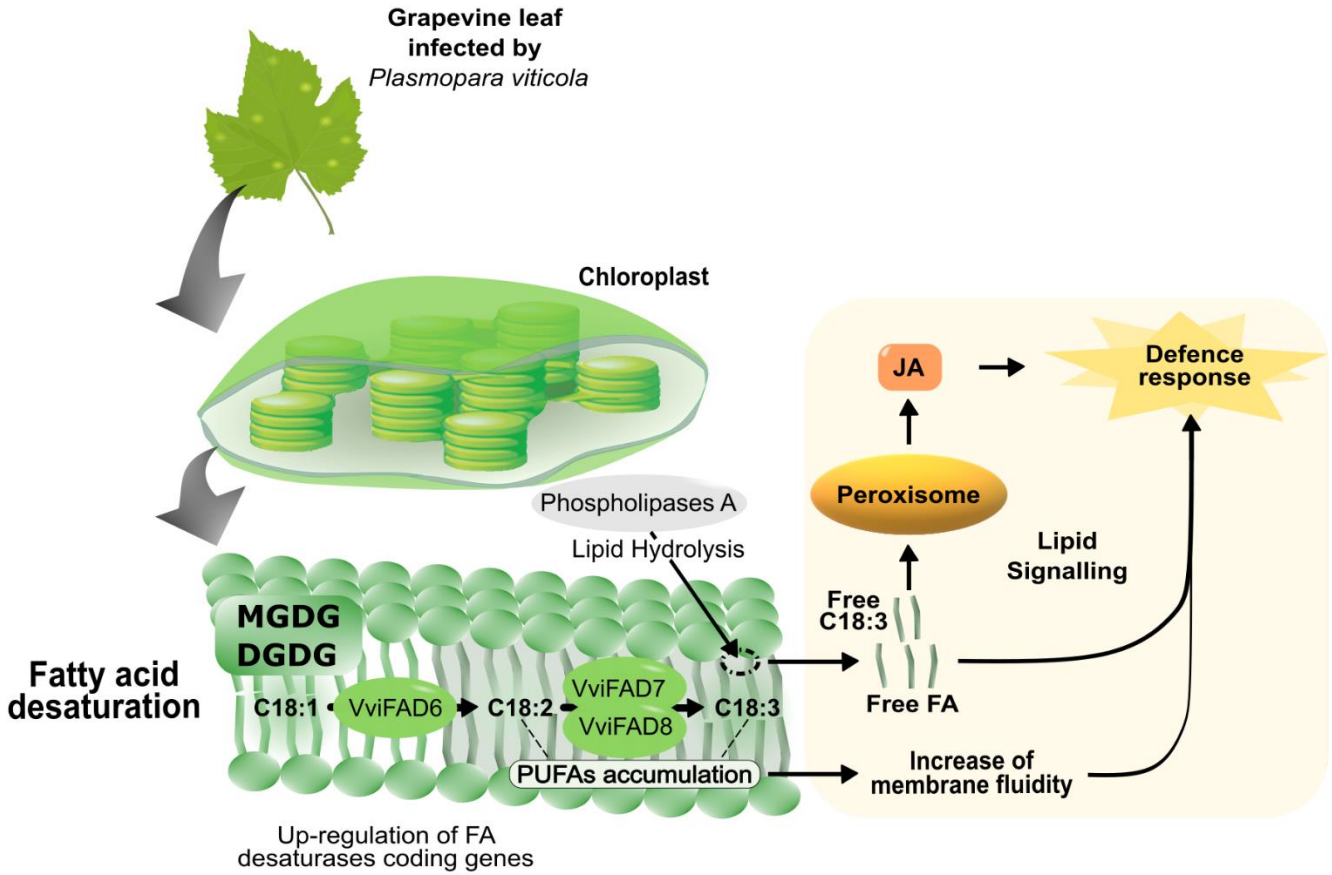
Increase Free FA pool

Activation of chloroplastidial *Fatty Acid Desaturases* (FAD6, FAD7, FAD8)



Increase FA desaturation

P. viticola inoculation triggers FA and lipid content alteration



Increase of membrane fluidity leads to the protection of photosynthetic apparatus

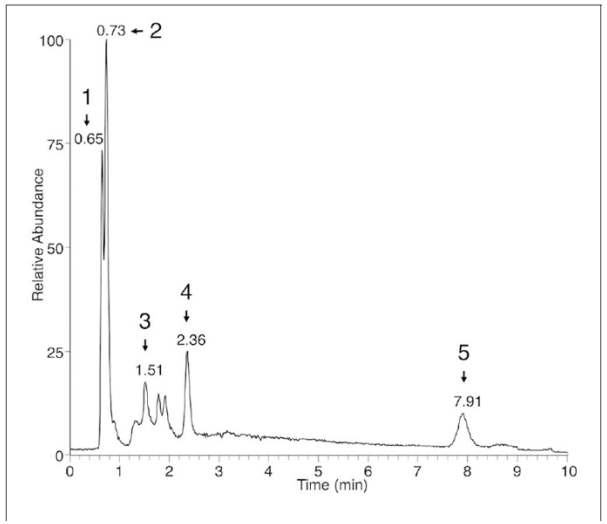
Higher availability of bounded and free C18:3, the Jasmonic Acid precursor

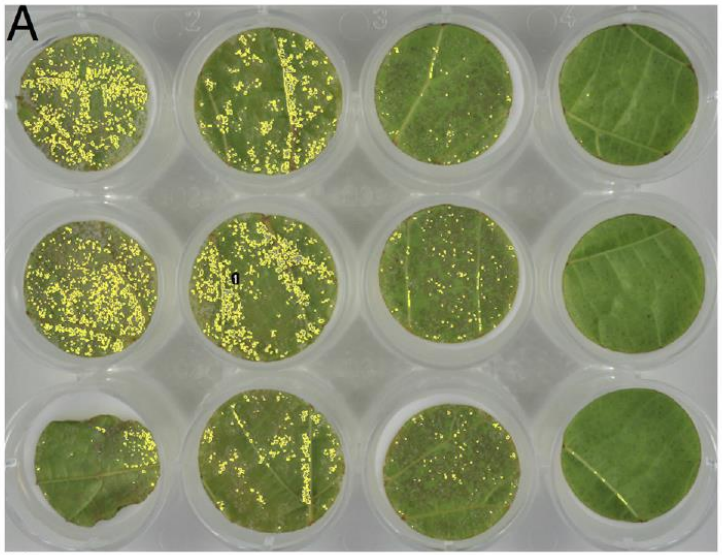
lipid signaling activation

Laureano et al. (2018) Sci. Reports
Laureano et al. (2021) *IJMS*
Cavaco et al. (2021) PPB

How about *Plasmopara viticola* lipids?

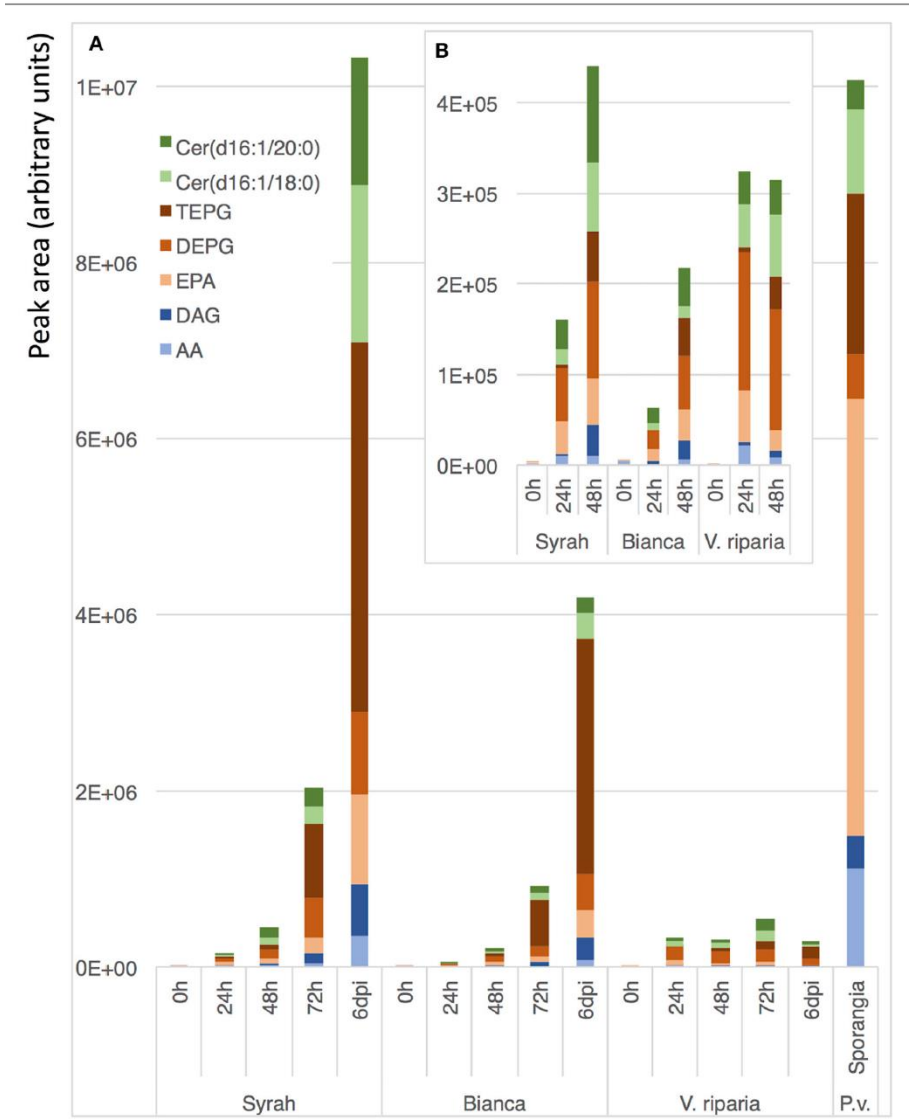
Are there pathogen specific lipids accumulated during the interaction?





B

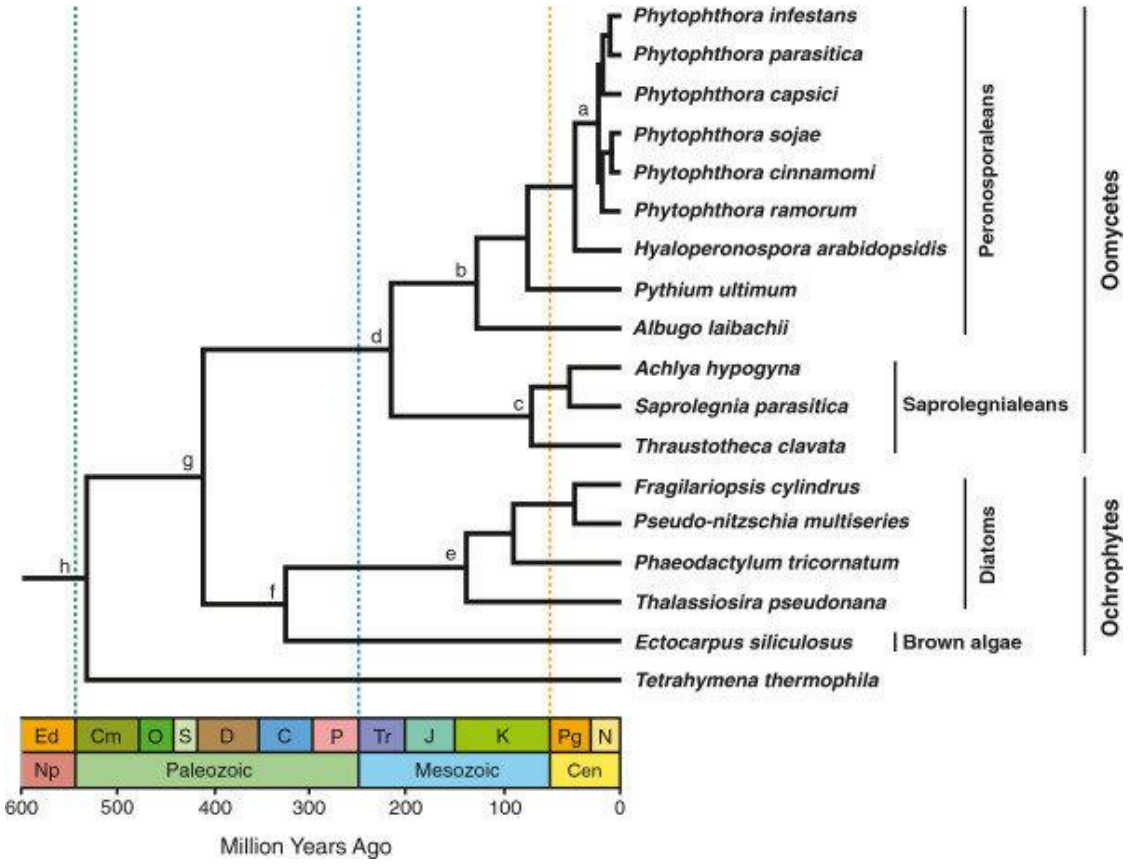
Metabolite	Pearson's correlation coefficient	P-value
EPA	0.849	1.37 e-18
AA	0.848	1.91 e-18
Cer(d16:1/18:0)	0.840	7.56 e-18
Cer(d16:1/22:0)	0.775	8.82 e-14
DAG	0.741	3.71 e-12
Cer(d16:1/16:0)	0.728	1.38 e-11
Cer(d16:1/20:0)	0.709	8.17 e-11
TEPG	0.702	1.47 e-10
AG	0.693	3.02 e-10
EPG	0.691	3.58 e-10
DEPG	0.689	4.18 e-10

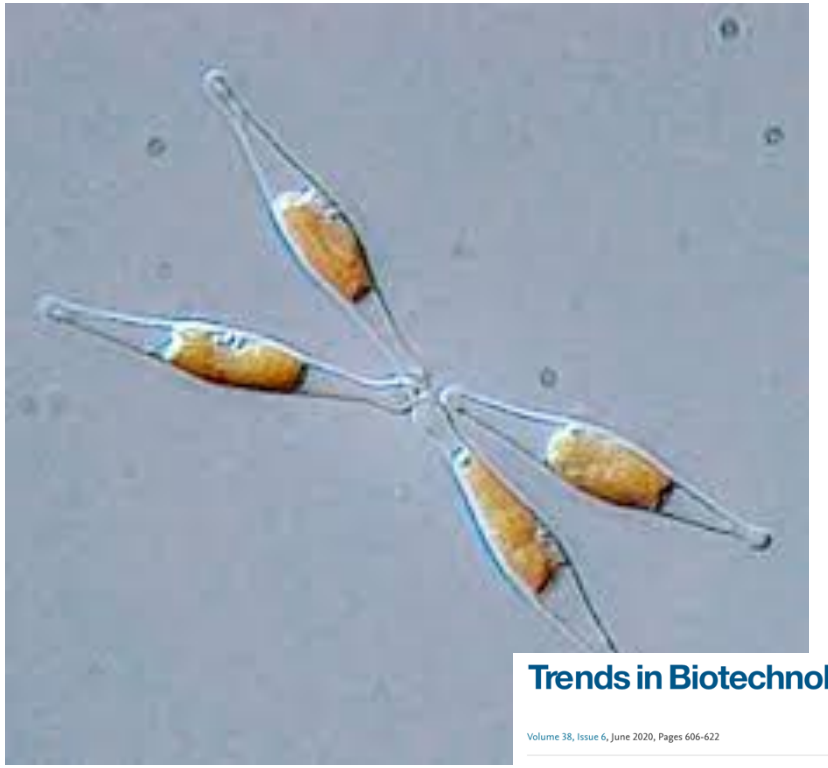


Negrel et al., 2018, Frontiers in Plant Science

How about *Plasmopara viticola* lipids?

Plasmopara viticola lipids that accumulate during infection and sporulation- do not exist in grapevine but there are a shared trait with diatoms.





Trends in Biotechnology

CellPress

Volume 38, Issue 6, June 2020, Pages 606-622

Review

Phaeodactylum tricornutum: A Diatom Cell Factory

Thomas Butler [1], Rahul Vijay Kapoore [1 2], Seetharaman Vaidyanathan [1]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.tibtech.2019.12.023 Get rights and content

Challenge

To define an experimental design

"My team has created a very innovative solution, but we're still looking for a problem to go with it."