UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola d'Enginyeria Agroalimentària
i de Biosistemes de Barcelona

MOLECULAR AND PHYSIOLOGICAL STUDY OF LIGHT REGULATION OF PLANT DEVELOPMENT

Bachelor's thesis

Biological Systems Engineering

Author: Lucía García Martín-Doimeadios

Tutor: Irma Roig Villanova

External tutors: Jaime F. Martínez-García

Pedro Pastor Andreu

16 / November / 2020

Abstract

A plant can detect a possible situation of shade by perceiving changes in the ratio between Red light and Far Red light (R:FR ratio). When it detects it, a wide network of components is activated. If the plant is a shade-avoider, the action of these components triggers changes in its development including vertical growth promotion, branching inhibition and early flowering, among others. Altogether these responses are known as the Shade Avoidance Syndrome (SAS) and are detrimental for agriculture.

A research group of the Centre for Research in Agricultural Genomics (CRAG) found the temporal and spatial action of three important components of this network. LONGHYPOCOTYL IN FAR-RED1 (HFR1), ELONGATED HYPOCOTYL 5 (HY5) and Phytochrome A (phyA) proteins act in the hypocotyl of Arabidopsis thaliana, mainly during the first stages of seedling development. Once this was determined, it remained to be seen whether these three genes were expressed in the same location where the proteins they encode act.

This bachelor thesis is focused on the study of hypocotyl elongation as a SAS response. The ultimate objective is the analysis of the spatial distribution of the expression of HFR1, HY5 and PHYA in the different organs of the Arabidopsis seedling.

To perform these experiments, Arabidopsis thaliana mutant plants deficient in each of these components were transformed with constructs containing the promoter of the gene fused to the coding sequence of the gene, fused to a reporter gene. We employed three binary plasmids: pHFR1:HFR1-GFP-3xHA (named pSP125), pHY5:HY5-YFP (named pB71) and pPHYA:PHYA-GFP (named pAPAG).

After that, we followed a pipeline to identify homozygous lines expressing the transgenes. This pipeline involved selection of transformants by identifying resistant lines to selective antibiotics or herbicides, molecular analyses via PCR and complementation tests.

It was possible to obtain segregating transgenic lines that showed partial complementation phenotypes for the pSP125 and the pAPAG constructs. Due to the situation of COVID-19 and some failed results, it was not possible to analyse the expression of the three genes in the seedlings.

Resumen

Una planta puede detectar una posible situación de sombra percibiendo los cambios en la relación entre la luz roja y la luz roja lejana (razón R:FR). Al detectarla, se activa una amplia red de componentes. Si la planta evita la sombra, la acción de estos componentes desencadena cambios en su desarrollo, incluyendo la promoción del crecimiento vertical, la inhibición de la ramificación y la floración temprana, entre otros. Estas respuestas se conocen como el síndrome de huida de la sombra (SAS; de sus siglas en inglés), y son perjudiciales para la agricultura.

Un grupo de investigación del Centre for Research in Agricultural Genomics (CRAG) estudió la acción temporal y espacial de tres importantes componentes de esta red. Las proteínas LONGHYPOCOTYL IN FAR-RED1 (HFR1), ELONGATED HYPOCOTYL 5 (HY5) y Phytochrome A (phyA) actúan en el hipocótilo de la *Arabidopsis thaliana*, principalmente durante las primeras etapas del desarrollo de las plántulas. Una vez determinado esto, quedaba por ver si estos tres genes se expresaban en el mismo lugar donde actúan las proteínas que codifican.

Este trabajo de fin de grado se centra en el estudio del alargamiento del hipocótilo como respuesta del SAS. El objetivo final es el análisis de la distribución espacial de la expresión de *HFR1, HY5* y *PHYA* en los diferentes órganos de la plántula de Arabidopsis.

Para realizar estos experimentos, las plantas mutantes de *Arabidopsis thaliana* deficientes en cada uno de estos componentes se transformaron con construcciones que contenían el promotor del gen fusionado a la secuencia codificante del gen, fusionado a un gen reportero. Empleamos tres plásmidos binarios: *pHFR1:HFR1-GFP-3xHA* (llamado pSP125), *pHY5:HY5-YFP* (llamado pB71) y *pPHYA:PHYA-GFP* (llamado pAPAG).

Después seguimos un proceso para identificar las líneas homocigotas que expresan los transgenes. Este proceso implicaba la selección de transformantes mediante la identificación de líneas resistentes a antibióticos o herbicidas selectivos, análisis moleculares mediante PCR y pruebas de complementación.

Fue posible obtener líneas transgénicas segregantes que mostraban fenotipos de complementación parcial para las construcciones pSP125 y pAPAG. Debido a la situación de COVID-19 y a algunos resultados fallidos, no fue posible analizar la expresión de los tres genes en las plántulas.

Resum

Una planta pot detectar una possible situació d'ombra percebent els canvis en la relació entre la llum vermella i la llum vermella llunyana (raó R: FR). Al detectar-la, s'activa una àmplia xarxa de components. Si la planta evita l'ombra, l'acció d'aquests components desencadena canvis en el seu desenvolupament, incloent la promoció del creixement vertical, la inhibició de la ramificació i la floració primerenca. Aquestes respostes es coneixen com la síndrome de fugida de l'ombra (SAS, de les seves sigles en anglès), i són perjudicials per a l'agricultura.

Un grup de recerca del Centre for Research in Agricultural Genomics (CRAG) va trobar l'acció temporal i espacial de tres importants components d'aquesta xarxa. Les proteïnes LONGHYPOCOTYL IN FAR-RED1 (HFR1), ELONGATED HYPOCOTYL 5 (HY5) i Phytochrome A (phyA) actuen en l'hipocòtil d' *Arabidopsis thaliana*, principalment durant les primeres etapes de el desenvolupament de les plàntules. Un cop determinat això, quedava per veure si aquests tres gens s'expressaven en el mateix lloc on actuen les proteïnes que codifiquen.

Aquest treball de fi de grau es centra en l'estudi de l'allargament de l'hipocòtil com a resposta del SAS. L'objectiu final és l'anàlisi de la distribució espacial de l'expressió de *HFR1, HY5* i *PHYA* en els diferents òrgans de la plàntula d'Arabidopsis.

Per realitzar aquests experiments, les plantes mutants d'*Arabidopsis thaliana* deficients en cada un d'aquests components es van transformar amb construccions que contenien el promotor del gen fusionat a la seqüència codificadora del gen, fusionat a un gen reporter. Fem servir 3 plasmidis binaris: *pHFR1: HFR1-GFP-3xHA* (anomenat pSP125), *pHY5: HY5-YFP* (anomenat pB71) i *pPHYA: PHYA-GFP* (anomenat pAPAG).

Després seguim un procés per identificar les línies homozigotes que expressen els transgens. Aquest procés implicava la selecció de transformants mitjançant la identificació de línies resistents a antibiòtics o herbicides selectius, anàlisi moleculars mitjançant PCR i proves de complementació.

Va ser possible obtenir línies transgèniques segregants que mostraven fenotips de complementació parcial per a les construccions pSP125 i pAPAG. A causa de la situació de COVID-19 i alguns resultats fallits, no va ser possible analitzar l'expressió dels tres gens en les plàntules.

Summary

Index of figures

Index of tables

Symbols and acronyms

Amp	Antibiotic used to prevent bacterial infections. Also used as a selective medium in experimental science. The acronym refers to Ampicillin.
bp	A measure of the length of double-stranded nucleic acids, usually DNA. The acronym refers to base pairs.
Cb	Bacterial antibiotic similar to Amp but more stable. The acronym refers to Carbenicillin.
CRAG	Centre for Research in Agricultural Genomics.
FR	Far red light (720-740 nm).
Gen	Antibiotic used in experimental science as an antibacterial used in cell culture to prevent contamination. The acronym refers to Gentamicin.
GFP	Green Fluorescent Protein. Protein that emits bioluminescence when excited by blue light.
HFR1	Gene called “Long Hypocotyl in Far-Red Light”. Directly regulated by PIFs.
Hg	Hygromycin. Antibiotic used in this case as a selective medium.
HY5	Gene called “Elongated hypocotyl 5”. Transcription factor that promotes photomorphogenesis in light.
Kn	Antibiotic used to treat several bacterial infections. Also used as a selective medium for some experiments. The acronym refers to Kanamycin.
pAPAG	Plasmid composed by the *PHYA* promoter fused to *PHYA*, fused to *GFP*.
pB71	Plasmid composed by the promoter of *HY5* fused to *HY5* gene, fused to YFP.
PCR	Polymerase Chain Reaction by its acronym, is a qualitative molecular biological technique for the detection of a fragment of genetic material.
PHYA	Phytochrome A protein. The active form of the protein is written phyA.
PIFs	Proteins that are transcription factors that bind to the active form of the phytochromes in the nucleus. Phytochrome-interacting factor by its acronym.

PPT	Chemical compound used as a pesticide for plants known as Phosphinothricin.
pSP125	Plasmid composed by the *HFR1* promoter fused to the GFP, fused to *HFR1* gene fused to a 3 histidine tag.
R	Red light (660-680 nm).
Rif	Antibiotic. Also used as an inhibitor for bacterial RNA polymerase. The acronym refers to Rifampicin.
SAS	shade Avoidance Syndrome by its acronym.
YFP	Yellow Fluorescent Protein. Usually used as a reporter gene.

Acknowledgements

I would like to thank Dr. *Irma Roig-Villanova* for offering me this opportunity, as well as *Pedro Pastor Andreu* and Dr. *Jaime F. Martínez García* for their expert advice and encouragement throughout this project.

1 Introduction

1.1 Background

This thesis is part of the research project on plant development and signal transduction regulated by light.

This research is led by Dr. *Jaime F. Martínez-García* at the *Centre for Research in Agricultural Genomics* (CRAG).

1.2 The importance of light for the plants: The Shade Avoidance Syndrome (SAS)

Plants usually grow in a highly changing environment. They have to compete for resources such as nutrients, water and also light. Light is an essential element for plants. It has a dual role: besides being an energy source for the photosynthesis, light is an information signal about the surrounding environment. The regulation of plant development by light is called photomorphogenesis (Franklin & Whitelam, 2007) .

One of the roles of the light as a signal is to inform about plant density. When plants grow very close together, they are more likely to shade each other. To this situation of shade, plants have different ways to react: either tolerating or avoiding it. If plants avoid shade, a competitive situation among them to capture more light is generated. Therefore, this competition influences plant development.

1.2.1 Plant photoreceptors and the Shade Avoidance Syndrome (SAS)

Plants have developed a series of photoreceptors that are able to detect different types of light. While carotenoids and chlorophylls are involved in the photosynthesis, other photoreceptors are involved in the photomorphogenic responses of the plant. Cryptochromes and phototropins are two types of photoreceptors which detect from UV-A (320-400 nm) to blue light (400-525 nm) (Figure 1). These types of lights induce responses such as phototropism and the opening of the stomatal pores in the leaf and stem epidermis, among others (Christie, 2007).

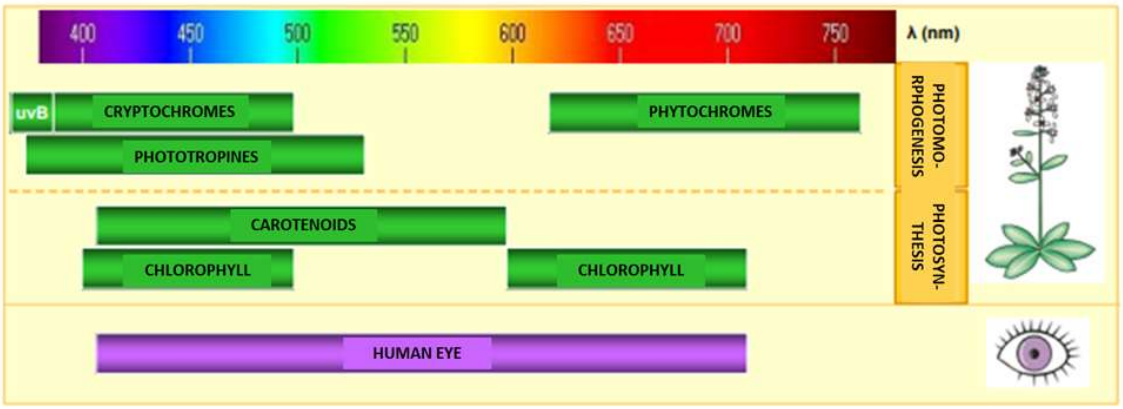

Figure 1. Representation of the different types of plant photoreceptors and the different wavelengths of light that they perceive, along with the main processes in which they are involved. Figure adapted from Roig-Villanova, 2007.

Red light (R, 580-650nm) and far-red light (FR, 680-740 nm) are detected by the phytochromes (Figure 1). This type of photoreceptors will be widely explained further on in section 1.2.3. The control seed germination, bulb formation, root development, dormant periods and the appearance of fruits and flowers (among others) is influenced by R (Runkle, 2016b).

In nature, plants rarely perceive monochromatic light, but a complex light signal that can vary in quality and quantity. From this, leaves mostly absorb blue and red light but reflect or transmit most far-red (Runkle, 2016a). The proportion or ratio of R and FR (that is, the R:FR ratio) that a plant receives informs about the presence of proximal vegetation. A high R:FR is perceived when the plant is growing alone or in low density. R:FR ratio decreases because the neighbouring plants mainly absorb red light whereas FR light is transmitted (and reflected). Therefore, a change in light quality (this decrease in R:FR) informs the plant about the presence of neighbouring vegetation and alerts it of a possible shaded situation in the future, before actually being in shade. In this situation there are no changes in the photosynthetically active radiation (PAR) that the plant perceives (Figure 2) (Roig-Villanova & Martínez-García, 2016).

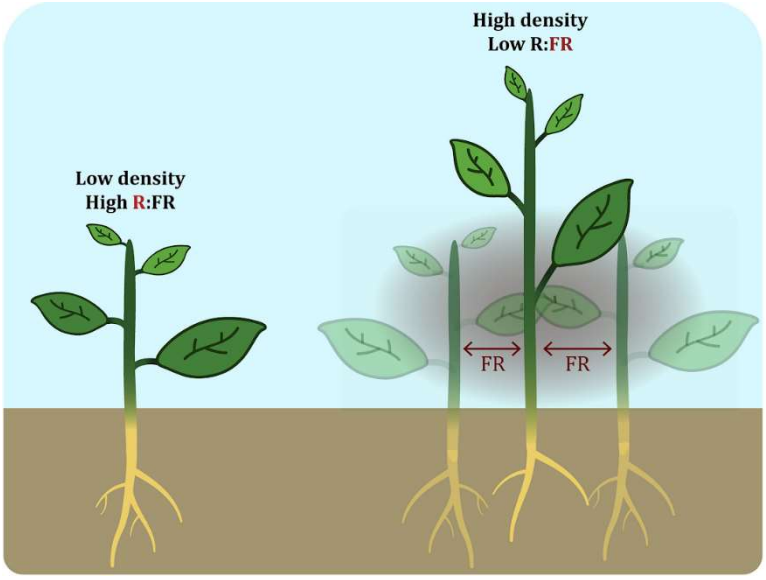

Figure 2. Representation of the relation between Red light (R) and Far red light (FR) depending on the vegetation density. Figure from Courbier & Pierik, 2019.

In response to the perception of the low R:FR by the phytochromes, a network of components comes into play promoting photomorphogenic responses such as inhibition of germination, early flowering, and vertical growth in an attempt to anticipate future competition for light. This phenomenon is called the "shade avoidance syndrome" (SAS) (Casal, 2012). In other words, when a shade-avoider species detects a reduction in the R:FR ratio, it activates responses that result in a redirection of resources, investing them in the growing of internodes and petioles in an effort to optimize light uptake, to the detriment of the growth of leaves and storage organs. As the SAS responses have a major negative impact in agriculture, a challenge for research is to understand these responses and to find the key components that regulate it.

1.2.2 *Arabidopsis thaliana* as a model plant to study the SAS responses

Arabidopsis thaliana is an annual plant from the mustard family (Brassicaceae or Cruciferae) widely used as a model in science. Its genome was the first plant genome to be completely sequenced. The reasons which make Arabidopsis a suitable model is its short generation time (approximately 2 months), self-pollinating and small genome, among others (Koornneef & Scheres, 2001). It has first a juvenile or seedling stage followed by a rosette habit during its vegetative growth, that could last

up to 39 days. Upon floral transition, it enters its reproductive growth by developing a main stem from the middle of the rosette. At the shoot apex of the main stem and in its branches the floral inflorescences develop, that upon fertilization will develop into the siliques that bear the seeds (Figure 3).

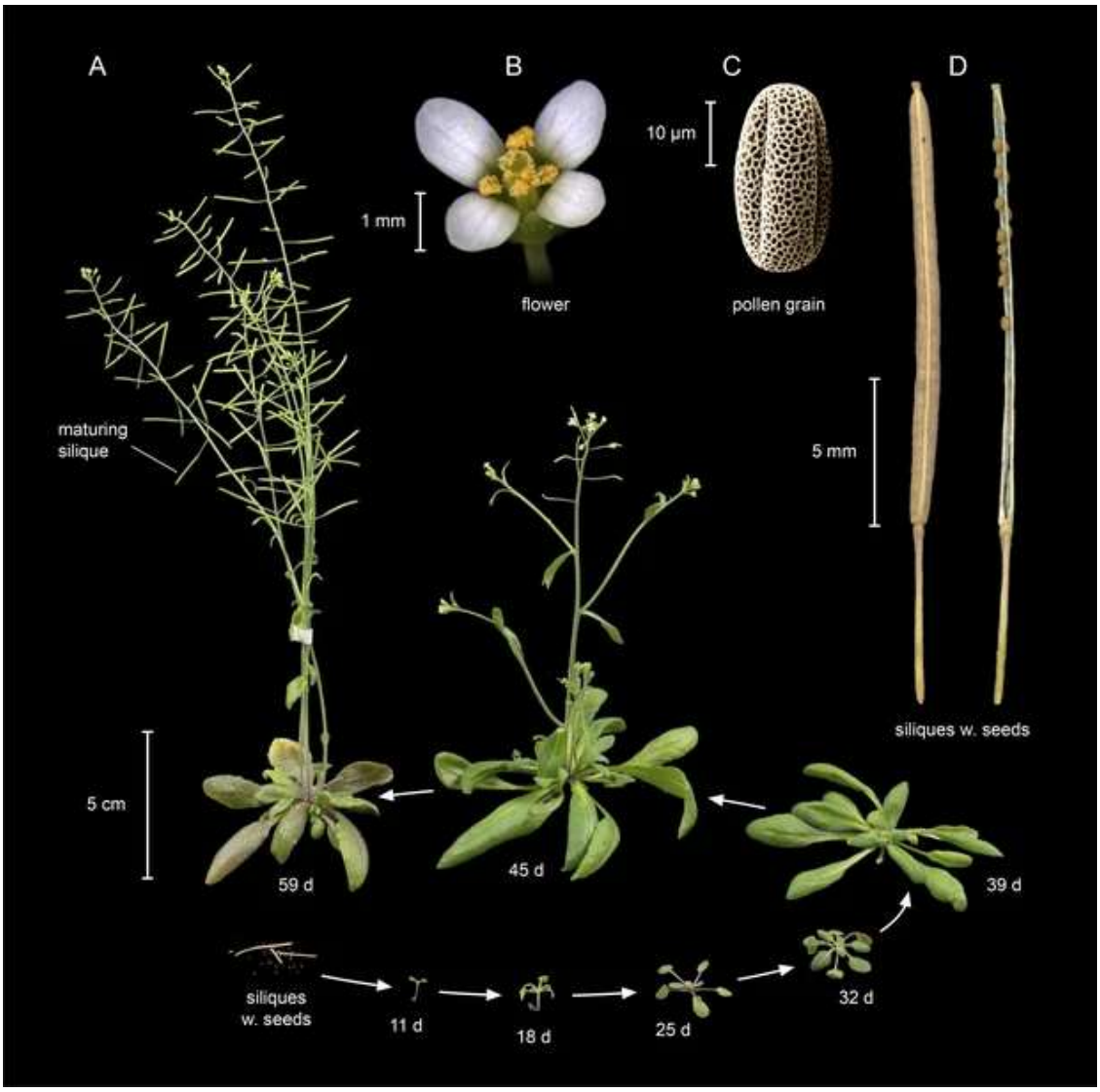

Figure 3. Life cycle of *Arabidopsis thaliana*. (A) The different development stages of *A. thaliana*. (B) The flower. (C) A pollen grains. (D) Mature siliques. Figure from Krämer, 2015.

Arabidopsis is a shade-avoider species, thus, it suffers SAS. In this plant the SAS includes responses such as repression of seed germination, promotion of the hypocotyl and stem elongation and inhibition of branching (Casal, 2012).

1.2.3 The phytochromes in *Arabidopsis thaliana*

In Arabidopsis there are 5 genes that encode for phytochromes, from *PHYA* to *PHYE*. The PHY proteins are synthetized in an inactive form and change their conformation and activity upon light exposure. Phytochromes have a molecule called chromophore that is in charge of perceiving light. Once the phytochrome protein absorbs red light, it is photoconverted from its inactive conformation (Pr) to its active conformation (Pfr). This active conformation induces a series of morphogenetic responses in the plant. On the other hand, when the phytochrome absorbs far red light, a photoconversion from the Pfr form to the Pr form takes place. In low plant density (high R:FR) the photoequilibrium of the phytochromes will also be displaced towards the active form, while in high plant density (low R:FR) it will be displayed towards the inactive form (Figure 4).

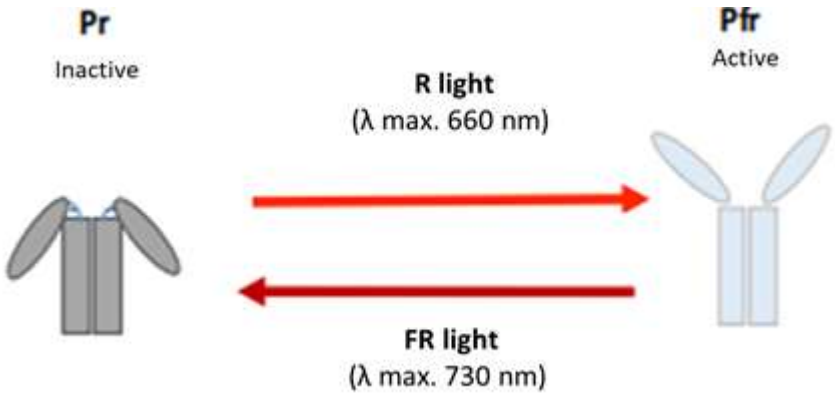

Figure 4. Diagram of the photoconversion of the inactive (Pr) and active (Pfr)phytochrome forms upon absorption of red or far red light. Figure adapted from Molina Contreras, 2017.

The phytochrome phyA is photo-labile. Its active form is more unstable than Pr. In those plants that are under strong shade conditions (low and very low R:FR), phyA accumulates, but as soon as light conditions change (high R:FR), phyA levels decrease rapidly. On the other hand, phyB to E are photostable.

Several experiments using mutants have shown that the most important phytochromes in regulating the SAS responses are phyA and phyB (Figure 5d). In Figure 5 it can be seen a summary of the action of phyA and phyB in the regulation of the hypocotyl response to shade.

In the situation of low density (high R:FR), phyA is photodegraded and the active form of phyB inhibits hypocotyl elongation (as there is no competition, the wild-type seedling does not need to elongate, fig 5a, d). In the situation of plant proximity, phyA is still photodegraded and phyB is in part inactivated. Therefore, the inhibition of hypocotyl elongation is not so strong, and the wild-type seedling elongates (figure 5b, d). When there is even more shade, for instance under a canopy, phyA is not photodegraded and together with phyB inhibits hypocotyl elongation (fig 5c, d). This can be interpreted in the following way: in a very shaded situation such as a canopy, even if the seedling elongates it will never reach the light, so it is not worth elongating (Martínez-García et al., 2014).

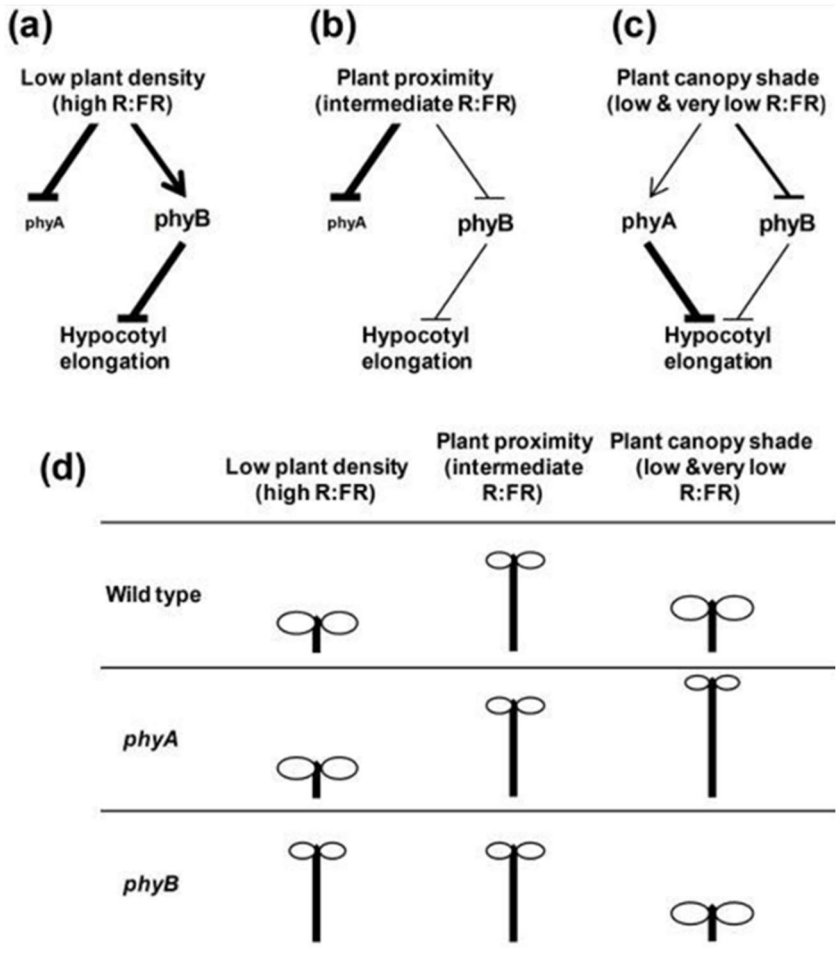

Figure 5. Hypocotyl elongation response of *phyA*, *phyB* and wild-type seedlings to different types of R:FR. Figure from Martínez-García et al., 2014.

1.2.4 Genetic components acting downstream phytochromes

Several molecular and genetic analyses have identified a series of components acting downstream the phytochrome photoreceptors with a role in the regulation of the SAS physiological and developmental responses (Roig-Villanova, 2007). In the next paragraphs I will describe the most important and/or relevant for this work (Figure 6).

The PHYTOCHROME-INTERACTING FACTORs (PIFs) are proteins which bind to the active form of phytochromes. When bound, PIFs are phosphorylated and later on degraded. Conversely, when the levels of the Pfr form of phyB are reduced under low R:FR, the PIF proteins are increased and the effect of SAS is enhanced (Casal, 2012). As the PIF proteins increase and accumulate, the active form of phyB is further reduced. The factors which have a greater effect in SAS are PIF4, PIF5 and PIF7. They are transcription factors responsible for regulating genes involved in cell expansion, resulting in growth and elongation of the hypocotyl (Casal, 2012). Altogether, they are considered positive regulators of the SAS responses.

HY5 is another component that can be considered a suppressor of the SAS. When the plant is in low R:FR ratios the expression of *HY5* is promoted. It is an important positive regulator of photomorphogenesis (Oyama et al., 1997). This component participates in ethylene-promoted hypocotyl elongation in the light but not in the dark (Yu et al., 2013). It binds to the promoters of cell elongation-related genes to regulate hypocotyl growth. HY5 can be considered as a PIF antagonist. While PIF enhances the elongation of hypocotyl, HY5 inhibits this effect.

HFR1 is also a negative SAS regulator, that counteracts the action of PIFs. That is, in low R:FR, HFR1 forms heterodimers with PIF4 and PIF5 and it binds to PIF7 (Zhang et al., 2019). By forming these heterodimers or bonds, PIF4, PIF5 and PIF7 are prevented from binding to the DNA and thus from activating downstream genes and enhancing the SAS (Fairchild et al., 2000).

Auxins are hormones that act in the regulation of plant growth. The synthesis of these phytohormones plays an important role in the elongation of the hypocotyl under shaded conditions. SAV3 promotes the biosynthesis of auxins and ethylene. The gene encoding this protein is expressed in the cotyledons. In the hypocotyl, this expression of *SAV3* is almost absent (Casal, 2012). The low ratio of R:FR strengthens the auxins signalling (in the cotyledons), but the expression

of *SAV3* is not increased by this ratio. SAV3 catalyses the formation of indole-3-pyruvic acid (IPA) from L-tryptophan in the auxin biosynthesis pathway (Casal, 2012).

YUCCA is another gene that encodes an enzyme involved in auxin synthesis. While SAV3 catalyzes the reaction of tryptophan to isopropyl alcohol, YUC (YUCCA) catalyzes the reaction of isopropyl alcohol to indoleacetic acid, which is an auxin (Casal, 2012).

PIF7 is directly involved in the regulation of the YUC enzyme. This PIF increases the activity of YUCCAs under a low R:FR ratio, causing more auxins to be synthesized.

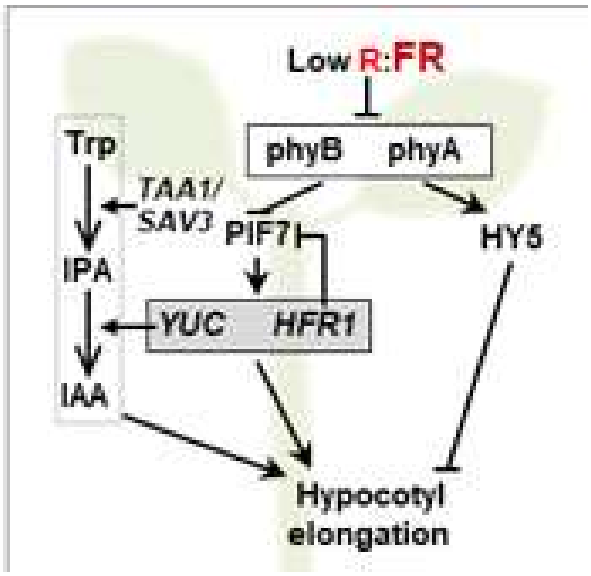

Figure 6. Diagram of the principal components involved in the regulation of the hypocotyl elongation in response to shade.

The research group of Dr. J. Martínez at CRAG found the temporal and spatial action of HFR1, HY5 and PHYA on the elongation of the hypocotyl. Pedro Pastor in his PhD discovered this through the use of mutants. He studied the effect that the mutation of each of these three components had on the "speed" of lengthening of the hypocotyl in response to simulated shade in days 2, 3, 4, 5 and 6 of seedling development and which cells of this organ were lengthened more (or less) at the end of the process. This data allows the determination of the time and the location of the HFR1, HY5 and PHYA proteins activity (data not published).

What remained to be seen is if, in order to exert that effect, *HFR1*, *HY5* and *PHYA* were expressed in the hypocotyl itself, in the cotyledons or everywhere in the seedling. This bachelor thesis has been put forward to provide more information on the expression of these components and to contribute to a better understanding of the regulation of SAS responses by the PHYA network through the generation of transgenic plants.

2 Objectives

The general aim of the project in which this bachelor thesis is included is to study the role of the proteins HY5, HFR1 and phyA in the response to shade of the *Arabidopsis thaliana* seedling. The laboratory of Dr. Martínez already discovered that these three proteins have differential action localizations and timings in the hypocotyl cells during its elongation in response to shade. What remains to be elucidated is if the genes encoding these proteins are expressed in the same organs in which the proteins act, or, if on the contrary, they are expressed in different organs and once the proteins are produced they move among cells.

To determine this, the main objective of this bachelor thesis was to analyse whether *HY5*, *HFR1* and *phyA* are expressed in the hypocotyl, the cotyledons or everywhere in the arabidopsis seedlings, and if this expression changes by simulated shade treatments. In order to achieve this, we initially planned to develop the following specific objectives:

1- Obtention of homozygous individual transgenic lines for three different plasmidic constructs producing HY5, HFR1 and phyA fused to reporter proteins in their corresponding mutant backgrounds.

2- Performance of complementation tests with the lines obtained in objective 1 in order to determine if the fusion proteins introduced are produced and functional.

3- Analysis of expression of *HY5*, *HFR1* and *PHYA* in *Arabidopsis thaliana* seedlings (hypocotyl and cotyledons) obtained in objective 1 by detecting the reporter proteins at the fluorescence microscope.

3 Materials and methods

3.1 Materials

3.1.1 Plasmidic Constructs

- pSP125 (*pHFR1:GFP-HFR1-3xHA*) - Construct that contains the promoter of *HFR1* gene (*pHFR1*) fused to the *GREEN FLUORESCENT PROTEIN* reporter gene (*GFP*), the *HFR1* gene (*HFR1*), and to 3 histidines tag (3xHA), in a binary vector with Phosphinothricin (PPT) resistance for selection in plants.

- pAPAG (*pPHYA:PHYA-GFP*)- Construct that contains the promoter of *PHYA* gene (*pPHYA*) fused to the *PHYA* gene (*PHYA*), and to the *GFP* reporter gene (*GFP*) in a binary vector with Kanamycin (Kn) resistance for selection in plants.

- pB71 (*HY5:HY5-YFP*)- Construct that contains the promoter of *HY5* gene (*pHY5*) fused to *HY5* gene (*HY5*), and to the *YELLOW FLUORESCENCE PROTEIN* reporter gene (*YFP*) in a binary vector with Hygromycin (Hg) resistance for selection in plants.

These constructs were previously obtained at the laboratory. They were used in order to generate homozygous transgenic lines for each of the different transgenes.

3.1.2 *Agrobacterium tumefaciens* strain

- C1C58/GV2260 - *Agrobacterium tumefaciens* strain for the cloning of binary vectors used for plant transformation. It is prepared to be transformed by electroporation. It has resistance to Amp.

3.1.3 Medium for *Agrobacterium tumefaciens* culture

In order for the transformed Agrobacterium strain to grow, a Yeast Extract Broth (YEB) medium (Vervliet et al., 1975) with a series of antibiotics was prepared. YEB medium: 5g/L beef extract, 1g/L yeast extract, 5g/L peptone, 5g/L sucrose, 2mL/L MgSO4 1M. pH 7,2. Sterilized by autoclave. The

working concentrations of the antibiotics used were 100µg/mL of Carbenicillin (Cb), 100µg/mL of Rifampicin (Rif) and 25µg/mL of Gentamicin (Gen).

3.1.4 Plant material

- Col-0 – Columbia 0 wild-type Arabidopsis thaliana ecotype.
- *hfr1-5* – Loss of function mutant for the *HFR1* gene.
- *phyA* – Loss of function mutant of the *PHYA* gene.

3.1.5 Plant medium

During these experiments, Arabidopsis was grown in different environments. For each environment, the growth medium/substrate changes:

- In the greenhouse the substrate for the Arabidopsis plants is a mixture composed by vermiculite, perlite, and peat in a proportion 1:1:1. The greenhouse conditions were: temperature of 22-24 °C during the day and 20 °C during the night. Relative humidity between 50 to 60 %. Long day (16 h light /8 h darkness) light conditions.

- The Arabidopsis seeds grow under *in vitro* conditions. The medium used is 0.5 % Murashige & Skoog medium including vitamins (Murashige & Skoog, 1962): 2,5g/L MS medium including vitamins, 0,25 g/L MES. 8g/L agar for the solid medium. pH 5,8. Sterilized by autoclave. Then, for the segregation analysis, different antibiotics are added to de MS medium. The antibiotic changes depending on the construct.
    - pB71: 30µg/mL of Hygromycin (Hg)
    - pAPAG: 25µg/mL of Kanamycin (Kn)
    - pSP125: 16µg/mL of Phosphinothricin (PPT)

3.1.6 Primers

In order to perform the molecular analysis, some primers are required to amplify the DNA sequence of interest. In the Polymerase Chain Reaction (PCR), primers are used as initiators. The DNA-Polymerase binds to the -OH group in order to start the synthesis process. The sequence of these primers depends on the construct as it has to be specific for it. For each DNA sequence to amplify,

a different pair of primers is needed. On Table 1 the different primers that have been used are specified.

Table 1. Specifications of the different primers used in PCR for the selection of the different constructs.

Construct	Primer		Sequence	
pB71	hy5-F	hy5-R	5'-CGGCTAGGTCTAAGAGATGG-3'	5'-GAACAGAGTTTCAGCTCAGC-3'
pAPAG	SPO40	SPO41	5'-GGCCATGGTAGATCTGACTAGTAA-3'	5'-GGCCATGGACACGTGGTGGTGGTGG-3'
pSP125	SPO02	SPO41	5'-TCATATGAAGCGGCACGACTT-3'	5'-GGCCATGGACACGTGGTGGTGGTGG-3'
	SPO40	SPO41	5'-GGCCATGGTAGATCTGACTAGTAA-3'	5'-GGCCATGGACACGTGGTGGTGGTGG-3'

3.2 Methods

3.2.1 Procedure to obtain transgenic plants

This process starts with Arabidopsis plants transformed (T_0) with *Agrobacterium tumefaciens* containing the construct of interest (Figure 7). A thousand seeds generated from the self-pollination of these plants (T_1 generation) are sowed in a selective medium. The seedlings that do not contain the transgene die, while the ones containing it (and therefore containing the gene of resistance for the selection) survive. A significant number of these surviving seedlings are transferred into soil and genotyped to confirm the presence of the transgene. They are considered independent transgenic lines, as they have the transgene incorporated in different places of their genomes. A hundred seeds generated from the self-pollination of each of these plants (T_2generation) are sowed in a selective medium. The segregation of the resistance will be analysed in these 100 seedlings, in order to determine the number of copies of the transgene that the line contains. The lines containing only 1 copy of the transgene (i.e., the ones that showed a proportion of 25% sensitive *versus* 75% of resistant seedlings) will be selected to continue with the study. 8 to 10 resistant seedlings from the selected lines are transferred into soil. 25 seeds generated from the self-pollination of each of these

plants (T_3 generation) are sowed in a selective medium. If all the 25 seedlings survive, this would mean that the T_2 generation plants already had 2 copies of the transgene (i.e. were homozygous), and therefore the T_3 generation is homozygous as well. Some of these seedlings are transferred into soil to obtain the T_4 generation and be kept as a homozygous stock for the transgenic lines.

The process to complete one generation Arabidopsis plants reaches 2-2 and a half month. Then, the whole process to obtain a homozygous transgenic plant takes approximately 1 year.

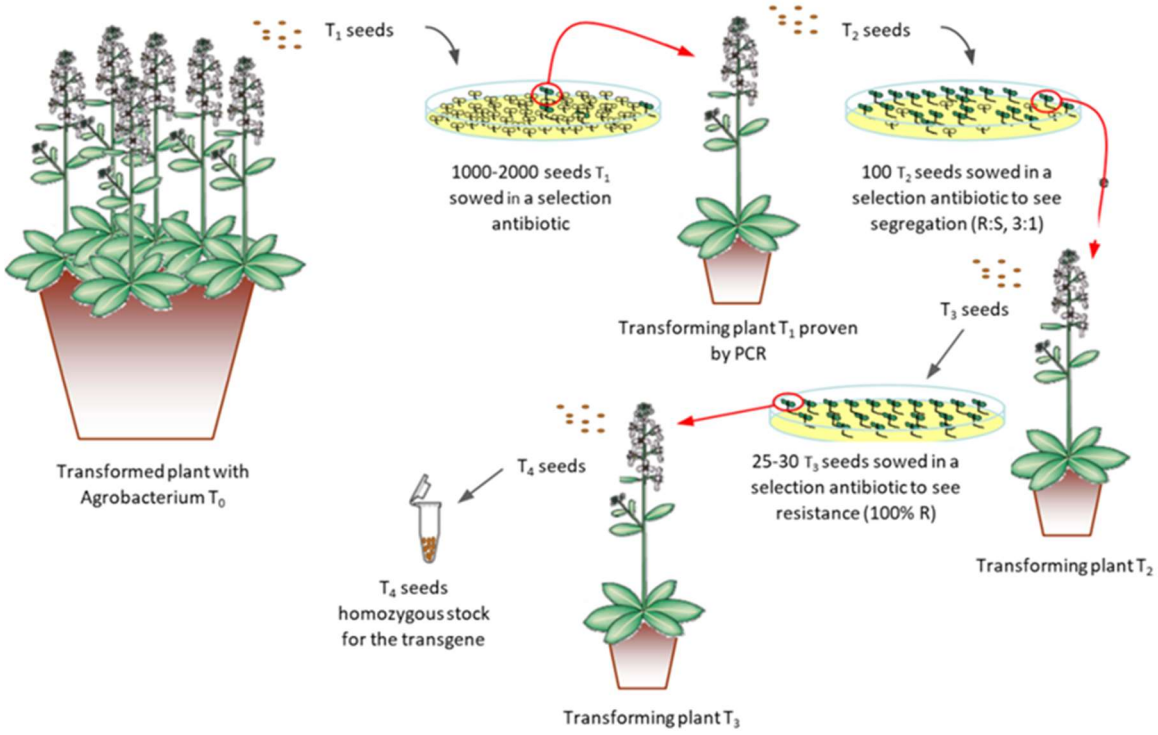

Figure 7. Schematic representation of the transformation of arabidopsis plants and selection of the different transforming generations. Figure from Roig-Villanova 2007.

In this work, I joined the process of obtention homozygous transgenic lines in different stages of the process described above, depending on the construct.

3.2.2 Transformation of *Agrobacterium tumefaciens*

Agrobacterium tumefaciens is used for the mutation of T_0 plant. In this case, I only performed the transformation of *A. tumefaciens* with the pB71 construct. The other samples were already transformed with pAPAG and pSP125 constructs.

A specific strain of *Agrobacterium tumefaciens* was selected (see section 3.1.2). The specific trait of this strain is that it is resistant to Amp. So that the transformed samples can grow in an Amp selective medium.

Electroporation was used as the technique for the transformation, so the strain selected is prepared for this technique. The process is the following:

1. Thaw a 50 µL aliquot of competent cells in ice for 5min (keep them on ice all the time).
2. Add 1 µL of 1:10 dilution in H_2O of plasmidic DNA, mix gently and incubate on ice for 5 min.
3. Transfer the mix into a chilled and sterilized electroporation cuvette. Keep it in ice (cuvettes are previously sterilized washing them with ethanol).
4. Electroporate the cells at 1.5kV (electroporator). Add immediately 950 µL of liquid YEB medium and transfer the mix to a new 1.5mL Eppendorf tube.
5. Incubate the cells 2 h at 28°C.
6. Plate 2 aliquots of 150 µL and 850 µL of the culture on plantes containing solid YEB medium.
7. Incubate the plate 2-3 days at 28°C upside down (to avoid condensation on the surface of the media) and wrapped in plastic (to avoid the medium dry).
8. Check the colonies by PCR.

3.2.3 Seed sowing for T_2 plants selection

The objective of T_2 selection is to obtain T_2 transformed plants, in order to continue the process and obtain T_3 seeds.

The segregation can be analysed thanks to the resistance of the construct to the antibiotic. For the pAPAG and pSP125 samples, T_2 seeds were already sowed. So, for pAPAG and pSP125 I started the

process by analysing segregation. For this, a count of germinated and non-sprouted seeds is made. The total number of seeds for T_2 is 100.

Once the count is done, the Chi-square statistical test tool is used. This statistical test is further explained in section 3.2.5.

Those plates that do not meet the conditions of the statistical test are discarded.

Once the plates with the selected seeds have been chosen, 8 germinated seeds are chosen per plate.

Each selected seed is sown in a pot with a 1:1:1 ratio of vermiculite, perlite, and peat, as indicated in section 3.1.5. During a week, the pots are kept covered with a plastic film to maintain high humidity. After this time, we considered that the seedlings had adapted to the greenhouse relative humidity and the film is removed.

After *Arabidopsis thaliana* has made floral transition and grown significantly, meaning that the main stem is around 14 cm, the plants have to be staked to avoid seed spread, as Figure 8 shows.

Figure 8. Example of staking a plant. In this figure you can see T_2 *Arabidopsis thaliana* transformed with pAPAG.

3.2.4 Statistical analysis of the resistance segregation

T_2 transformed plants will be considered as transformed if seeds germinate in a medium containing the selection agent. For example, if a pSP125 T_2 seed germinates in the presence of the PPT herbicide, it will mean that the seedling is resistant to it and therefore it contains the transgene. If we have a 75% of resistant plants, this will mean that they come from a T_1 plant with only one copy of the transgene, and therefore the T_2 resistant seedlings have at least one copy of it (see Figure 9).

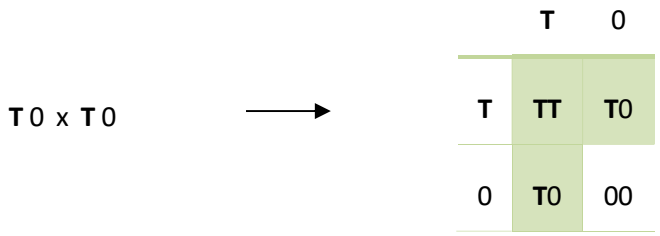

Figure 9 Representation of the genetic segregation of a T_2 offspring derived from a self-fertilized T_1 transformed plant, following the Mendelian laws of segregation. Where "T" is a symbol for a copy of the transgene and "0" is a symbol for no copy of the transgene.

As said before, ideally 75% of the T_2 seeds should be resistant to the antibiotic and then grow because they are transformed. However, when Mendel laws are applied to a real case the samples do not follow strictly these percentages. So, for the selection of the transformed seedlings, the Chi-Square statistical test is used. Chi-square test allows us to determine if the observed segregation of the transgenic lines that we are analysing follows the expected segregation (75% R/ 25% S).

Following the next equations, we can conclude statistically if the sample has incorporated the transgene or not.

$$\text{Chi-square}\left(X^2\right)=\frac{(\text{Observed}_r\text{-Expected}_r)^2}{\text{Expected}_r}+\frac{(\text{Observed}_s\text{-Expected}_s)^2}{\text{Expected}_s} \qquad \text{Equation 1}$$

The X^2 value has to be calculated following the Equation 1 either for the sensitive seedlings and for the resistant ones. For the observed value, the number of sensitive or resistant seedlings has to be counted for each sample in each plate. The expected value for the sensitive (s) seedlings

corresponds to the 25% of the total number of seedlings. On the other hand, the expected value for the resistant (r) seedlings correspond to the 75% of the total seedlings.

Once the Chi-square value is calculated, we define the degrees of freedom by the following equation.

$$\text{Degrees of freedom (DF)} = n-1 \qquad \text{Equation 2}$$

On the Equation 2 , n is the number of the types of plant, in this case corresponding to sensitive or resistant. Which means that the degrees of freedom are 1.
We consider that in order to reject the idea that this data could be a result from the fate, we assume a 95% of confidence. Meaning that if the result is greater to the 5%, then it will be considered that the result is due to the work of chance.

Table 2. Chi-square table for 1 to 4 degrees of freedom and with a 0.1, 0.05, 0.02, 0.01 significant level. Adapted from Fisher, R. A. (1954).

DF	0.1	0.05	0.02	0.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.345
4	7.779	9.488	11.668	13.277

If the calculated Chi-square value with the Equation 1 is lower than the Chi-square value of the table, then it will mean that the sample has inserted the transgene and therefore, this sample can continue to carry on the experiment. If the calculated Chi-square is greater, then the sample is excluded from the experiment. Once the seedlings were considered as transformed, the molecular analysis could be performed.

3.2.5 Genotyping of T_2 seedlings

The aim of this experiment is to confirm that the plants have been transformed with the different constructs.

This genotyping process should have been performed in T_1 seedlings in order to detect if the seedlings had been transformed. However, due to the COVID-19 lockdown, it was not possible to

work in the laboratory while T_1 seedlings were developing. Therefore, the genotyping had to be done on pools of T_2 seedlings.

The process was as follows:

3.2.5.1 Genomic DNA extraction

The Edwards protocol (Edwards et al., 1991) was used for the DNA extraction with a modification at the beginning of the process.

1. Deposit 6 to 8 germinated seeds in 1.5 mL *Eppendorfs* and add 2 glass beads.
2. Leave the *Eppendorfs* in an ice box.
3. Place the *Eppendorfs* in the Tissue-Lyser for 1 minute and 30 seconds with stirrings of 30 ms.
4. Remove the *Eppendorfs* from the Tissue-Lyser and put them back on ice.
5. Add 400 µL of extraction buffer composed of 200 mM Tris-HCl pH 8.0, 250 mM NaCl, 25 mM EDTA, 0.5% SDS.
6. Vortex 5 seconds; leave at room temp until all preps are ready.
7. Centrifuge 1 min in microfuge (13,000 RPM).
8. Transfer decanting one Eppendorf content into the other directly, 400 µl of supernatant to a new tube.
9. Add 400 µl of isopropanol, mix, leave at room temperature for 2 min.
10. Centrifuge 5 min in microfuge, decant supernatant.
11. Wash with 500 µl of 70%EtOH.
12. Centrifuge 5 min in microfuge, decant supernatant, air-dry pellet 10-15 min.
13. Resuspend in 50 µl of TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8).
14. Store at 4°C.

3.2.5.2 Polymerase Chain Reaction (PCR)

1. For the preparation of the PCR, the different components should be mixed (see Table 3). It is important to keep in consideration that apart from the samples to be examined, a negative and positive control must be included in the analysis.

Table 3 Concentrations of the different components of a standard PCR reaction

	Concentrations
DNA template	1 - 200 ng
dNTPs	200 µM
Forward primer	50 - 100 pmol
Reverse primer	50 - 100 pmol
Taq Polymerase	2.5 - 5 U
mQ-H_2O	Add to q.s. to 50 µL

2. For each specific PCR relation, program the thermal cycler making special attention to the annealing temperature (depending on the primers used, due to its melting temperature), and the extension time (determined by the size of the band that we want to amplify) . Table 4 is an example of the program for the Thermal cycler for the detection of the pSP125 construct by PCR.

Table 4. Thermal cycler program on genomic DNA for the detection pSP125 plasmid using the SPO02 and SPO41 primers.

	Time (s)	Temperature (°C)	
Initial denaturation	180	94	
Denaturation	30	94	35 cycles
Annealing primer	30	55	
Extension	35	72	
Final extension	180	72	
	HOLD	15	

3.2.5.3 Gel electrophoresis

This is the final step of the genotyping experiment.

1. Gels are made at 1% agarose in TBE 1X. (TBE 10X: Tris base 0,089 M, boric acid 0,089 M, EDTA 20mM. pH 8).
2. Heat the mixture in a microwave until agarose is completely dissolved.
3. Cool it down to approximately 60ºC.
4. Once the solution has cooled down, add ethidium bromide in an extraction hood.
5. Mix well and pour the solution it into the gel bed. Leave it to solidify.
6. Once solidified, introduce the gel in its bed in the electrophoresis tank containing TBE 1X.
7. Load 12 µL of each PCR sample into each well of the gel. Leaving the first lane free.
8. Load the molecular marker into the first and last lane.
9. Once the samples and the marker are loaded, run the gel at 110 kV
10. Visualize the result observing the gel in an ultraviolet light chamber.
11. Determine the size of your PCR product band by comparing it with the known-size bands of the molecular marker.

3.2.6 Complementation test of T_2 seedlings

Complementation test are normally performed in T_3 or T_4 seedlings because it is in this generations when the plants are homozygous for the transgene (Figure 7). If homozygous plants (this is, with 2 copies of the transgene), correctly expressed the transgenic proteins, they might be able to complement the mutation and plants should phenotypically behave as the Col-0 wild type. However, even if a plant has only one copy of the transgene (for instance, in the T_2 segregating population), it can be expressing the transgene and therefore could behave like the Col-0 wild-type as well. Taking this in mind, we performed the complementation test in T_2 seedlings. As it is a segregating population, in order to achieve some statistical similarity, more seeds were taken.

For this process, the following steps were done:

3.2.6.1 Sterilization of the seeds

1. Imbibe the seeds in 1.5 mL *Eppendorf* with 1 mL sterile water 0.1% w/v Tween 20 – first vortex a couple of times. Then incubate 20 minutes on the bench.
2. Remove the imbibition solution and replace it with 1 mL of sterilization solution. For this sterilization solution, mix 1 mL of commercial bleach, 100 µL of 10% (v/v) Tween 20 and 8.9 mL of MilliQ water. Incubate for a maximum of 10 minutes (from this point onwards work under the flow cabinet).
3. Remove sterilization solution and wash 5 times with a 1 mL of sterile water.
4. Keep the last 1 mL of sterile water on the *Eppendorf.*

3.2.6.2 Sowing the seeds

After sterilization the seeds, a sowing process is required for the complementation test. As said before, since these are T_2 seeds, in order to reach a statistical similarity, 50 seeds of each sample were taken.

1. Prepare sterile plates containing 0.5 x MS solid medium with vitamins.
2. After the medium is dry, divide each plate with a marker (on the outside of the plate) into two parts. Label every part with the corresponding sample.
3. With a 1 mL pipette, take the maximum number of seeds. Then, place it one by one into each part of the plate. Until having 50 seeds approximately. You can use a guide to spread the seeds with a uniform distance.

Figure 10. Approximately 100 T2 pSP125 seeds sowed for a complementation test in 0.5 MS with vitamins.

4. After sowing every sample, the result has to be similar to Figure 10. It must be considered that controls must be sewed. The controls of the mutant phenotype will be the mutant seeds (*phyA* seeds in the case of pAPAG lines, and *hfr1-5* seeds in the case of pSP125 lines), while the control of the wild-type phenotype will be Col-0 seeds.
5. After sowing, all plates are stratified, to synchronize their germination. For this, they will be kept for 2 days in darkness and cold (covered with aluminium foil in a fridge 4ºC).

3.2.6.3 Simulated shade treatments

After stratification, all plates are kept 2 days in white light (W) 22 °C. Thereafter, half of the plates are kept in W while the other half are kept 5 days in W +FR (Figure 11) This is the treatment that simulates shade in the laboratory. It consists in white light enriched with FR (with LED lamps), causing a low R:FR and therefore generating a simulated shadow situation.

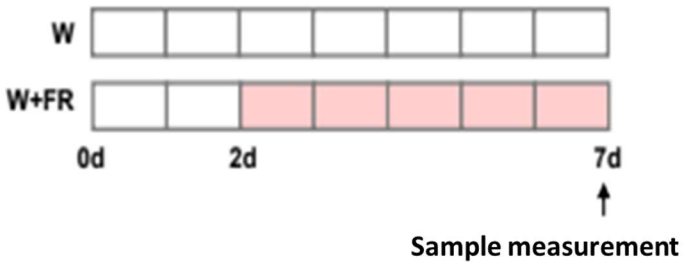

Figure 11 Representation of the simulated shade treatment for the phenotype analysis. Figure form Roig-Villanova 2007.

3.2.6.4 Hypocotyl measurement

After the 7 days of simulated shade treatment, hypocotyls are measured in order to compare the measurements of the transgenic lines to the wild-type Col-0 and mutant controls, and to determines if they behave as the wild-type and in consequence, complementation has been achieved.

Each seedling laid on the medium in the plate so that the entire hypocotyl can be seen.

After every seedling is laid down, the plates are placed next to a ruler and photographed.

Then, with the help of a program called “ImageJ”, it is possible to measure the hypocotyl elongation by establishing a scale between the ruler and the pixels of the image.

Hypocotyl is the part of the stem of a germinated seed that goes from the Start of the root to the start of the cotyledons. In Figure 12 the Arabidopsis hypocotyl is represented. In this case, Arabidopsis is a dicotyledon species, so it has two leaves emerging firstly from the embryo.

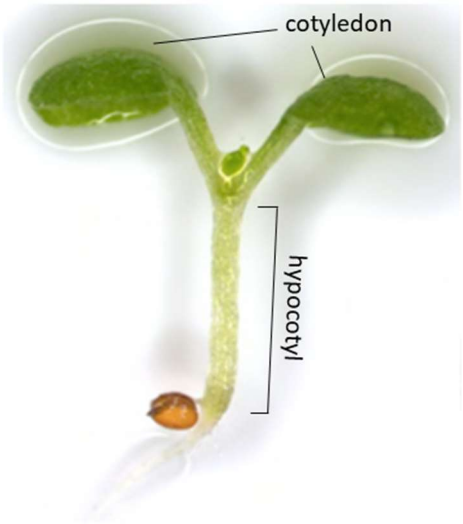

Figure 12 Photograph of an Arabidopsis seedling, where the hypocotyl and the cotyledons are shown. Figure adapted from Andersen et al., 2014.

After measuring all the hypocotyls, the genetic complementation is analysed with the help of the data collected by creating graphic diagrams such as a boxplot.

4 Results and discussion

4.1 Pipeline of the analysis

As it has been mentioned before, the final goal of the project was to analyse the expression patterns of *HFR1*, *HY5* and *PHYA* in the different organs (hypocotyl, cotyledon) of *Arabidopsis thaliana* seedlings in response to shade. To do that, firstly we had to generate transgenic plants containing reporter lines for these genes in its corresponding mutant background (Objective 1). The three constructs used in this work to generate the transgenic lines were previously obtained in my host laboratory using standard molecular biology cloning protocols. They all have in common the fusion of the promotor of the gene, the coding sequence of that gene and a reporter gene.

In order to verify that the fusion promotor + coding sequence of the gene replaces the function of the endogenous promotor and coding sequences, a complementation test must be performed (Objective 2). The complementation test consists in analysing the phenotype of the homozygous plants for the transgene in the mutant background in comparison with the original mutant and the wild type. If in the transformed mutants the wild-type phenotype is restored, this means that the fusion construct gives a functional protein. Thus, if this construct also contains a fluorescent protein, we can assume that the activity of this protein localizes where the endogenous gene is expressed, and we can use these lines for expression analysis (Objective 3).

In the next sections I will explain the advances that I obtained with the different constructs.

4.2 Obtention of homozygous seeds for the pB71 (*pHY5:HY5-YFP*) construct in the *hy5* mutant background

The construct that we used to analyse the expression of *HY5* contains the promoter of *HY5* (*pHY5*) fused to the coding sequence of *HY5* and to the *GFP* reporter gene. The name pB71 is an internal short code for this construct that the host laboratory uses to simplify the nomenclature.

4.2.1 Transformation of Agrobacterium with the pB71 construct

When I started working on this project, the construct to generate this transgenic line was already generated.

I tried to transform the construct into *Agrobacterium tumefaciens* via electroporation 4 times. On the last attempt, three *Agrobacterium* individual colonies had grown in selective medium. These colonies were candidates to contain the transgene and were chosen for further analysis (Figure 13).

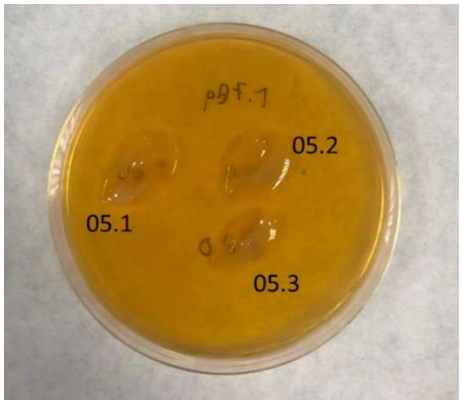

Figure 13. Three colonies of pB71 plasmid transformed with *Agrobacterium tumefaciens* using electroporation cells.

Figure 14 PCR of 05.1, 05.2 and 05.3 *Agrobacterium tumefaciens* colonies. using a 1:100 dilution of pB71 plasmid as positive control and H_2O as negative control. The primers used are hy5-R and hy5-F.

A genotyping analysis was performed to check the presence of the pB71 plasmid in the selected colonies. Therefore, a PCR analysis using specific primers to detect a 250 base pairs (bp) band of the *HY5* gene was performed. As a control we used a 1:100 dilution of the construct. As Figure 14 shows, the molecular analysis was negative in all three samples. With these results we might conclude that none of the colonies contained the transgene. However, colony PCRs from *A. tumefaciens* are tricky because of the cell wall of the bacteria. We could try to improve the success of the PCR by performing a DNA extraction of the colonies in order to obtain the construct and better detect it. Alternatively, we could transform again *A. tumefaciens* in order to obtain new colonies to screen. As we had other 2 constructs to follow, we stop here working on this one.

4.3 Obtention of homozygous seeds for the pAPAG (*pPHYA:PHYA-GFP*) construct in the *phyA* mutant background

The construct that we used to analyse the expression of *PHYA* contains the promoter of *PHYA* (*pPHYA*) fused to the coding sequence of *PHYA* and to the *GFP* reporter gene. The name pAPAG is an internal short code for this construct that the host laboratory uses to simplify the nomenclature.

4.3.1 Selection for pAPAG T_2 sowing

When I started working with this line the T_2 generation of transformed plants was already available in the lab. T_2 plants are genetically heterozygous. These plants have at least one copy of the transgene. Meaning that when they self-fertilize 25% of the offspring will be wild-type, while the other 75% will have one or both copies of the transgene, as Figure 9 shows.

Briefly, 100 T2 seeds from each independent line were sowed on MS + Kn media, and the number of resistant or sensitive seedlings was counted. Later, a Chi-square statistical test (widely explained in section 3.2.4) was performed to determine if they contained one or more copies of the transgene calculated.

14 T_2 lines were analysed (Table 5). All of them germinated correctly. From this, seven samples were considered to contain 1 copy of the transgene and were transferred into pots to obtain the following generation. The percentage of resistant seedlings in these seven samples is equal or greater than the 75% of the total amount of seeds, as the Chi-square conditions determine. Since the DF (Degrees of Freedom) and the level of confidence of the Chi-square test has not changed (section 3.2.4), the Chi-square value of the Table 2 is the same. And for the Chi-square calculated value, Equation 1 was used.

Table 5 Results of the selection of T_2 pAPAG plants. On the first column are all the 14 pAPAG T_2 plants. In green are the ones which are considered transformed with one copy of the transgene, based on the Chi-square statistical test ($\chi 2$ on the sixth column). In red are the ones dismissed from the experiment.

plants KnR	KnR	KnS	if $\chi 2<3.84$, 1 T-DNA	grow 8 kana^R seedlings to pots
APAG.01	93	10	12,84	NO
APAG.02	95	10	13,41	NO
APAG.03	80	26	0,01	YES
APAG.04	96	8	16,64	NO
APAG.05	85	16	4,52	NO
APAG.06	83	19	2,21	YES
APAG.07	97	4	23,84	NO
APAG.08	98	4	24,17	NO
APAG.09	82	26	0,05	YES
APAG.10	83	19	2,21	YES
APAG.11	85	22	1,2	YES
APAG.12	90	15	6,43	NO
APAG.13	78	24	0,12	YES
APAG.14	78	27	0,03	YES

As the Table 5 show, from this analysis we conclude that the lines pAPAG.03, pAPAG.06, pAPAG.09, pAPAG.10, pAPAG.11, pAPAG.13 and pAPAG.14 contain one single copy of the transgene and they are selected to proceed with the obtention of the T_3 generation.

4.3.2 Genotyping of pAPAG T_2 plants

As said in section 3.2.5, the molecular analysis should had been performed in T_1 samples. However due to the COVID-19 lockdown, it was not possible to perform the genotyping of this samples. So, the molecular analysis is done in pools of the T_2 samples.

Based on the results of the segregation analysis, the genotyping was done in samples pAPAG.03, pAPAG.06, pAPAG.09, pAPAG.10, pAPAG.11, pAPAG.13 and pAPAG.14 .

For the molecular analysis for pAPAG, the Polymerase Chain Reaction (PCR) was used. The primers used were specific to the reporter protein GFP, expected to amplify a band of approximately 760 pb. For the control samples, a pAPAG dilution (1:100) was used as a positive control, and for the negative control H_2O was used.

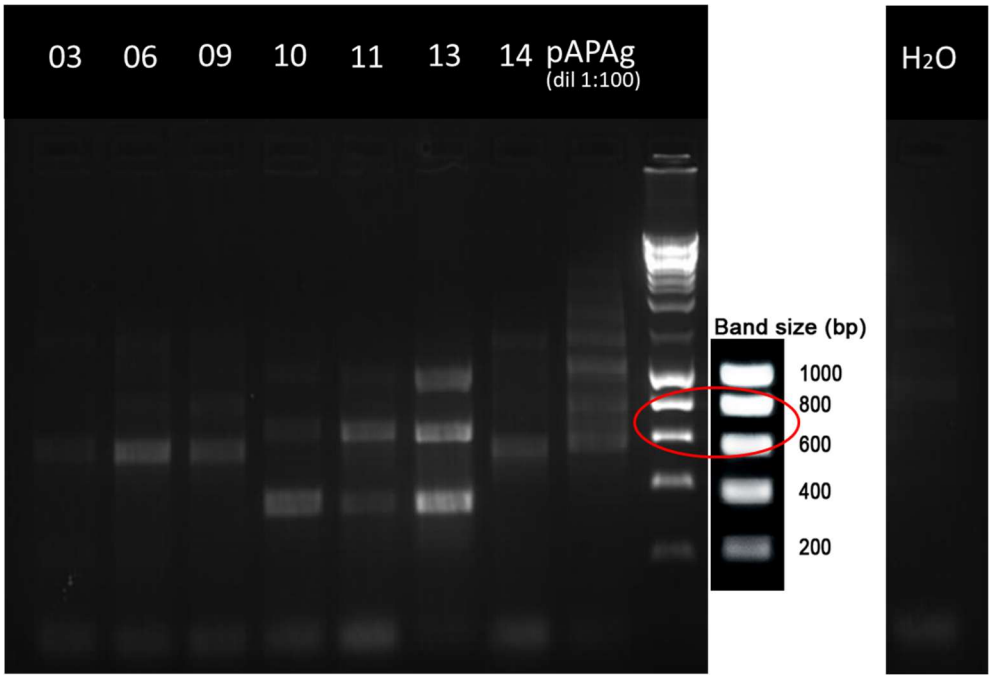

Figure 15. pAPAG PCR results of T_2 plant genotyping with SPO40 and SPO41 primers. 7 samples (pAPAG.03, pAPAG.06, pAPAG.09, pAPAG.10, pAPAG.11, pAPAG.13 and pAPAG.14) plus the positive and negative control (pAPAG diluted plasmid and H_2O respectively) are shown. Between the positive and negative control is the DNA ladder with the known bp of the bands.

As Figure 15 shows, these molecular analyses cannot be admitted as a valid result. The positive control did not work, as it does not have the expected band.

4.3.3 Complementation test of the T_2 pAPAG lines

As said before (section 3.2.6), the complementation test is normally performed in T_3 or T_4 samples. However, since resistant T_2 seedlings should have at least a copy of the transgene, the complementation test can be performed at this point of the process. In this section, the results of the complementation test in simulated shade and in normal conditions of T_2 pAPAG seedlings are represented. A simulated shade treatment is needed because under these conditions the mutant

samples and the Col-0 wild type have different phenotype. Therefore, it is the only condition in which it can be detected the complementation of the different lines.

The complementation test for the pAPAG T_2 samples under shade treatment had to be repeated twice because *phyA* did not grow as it was expected. So, the experiment was repeated with newer seeds.

In these boxplots are represented the hypocotyl measurements under simulated shade or normal conditions after seven days. In order to determine the differences between the simulated shade treatment and the light treatment, the complementation test in pAPAG was performed under W light conditions as a control test.

The results of the hypocotyl measurements under light conditions are shown in Figure 16. The samples seem to have a tendency where the majority of the measurements are from 1 mm to 2 mm. There are some samples missing, as there were not enough seeds to perform the experiment twice. So, for samples *phyA* and pAPAG.14 the complementation test under light treatment could not be performed.

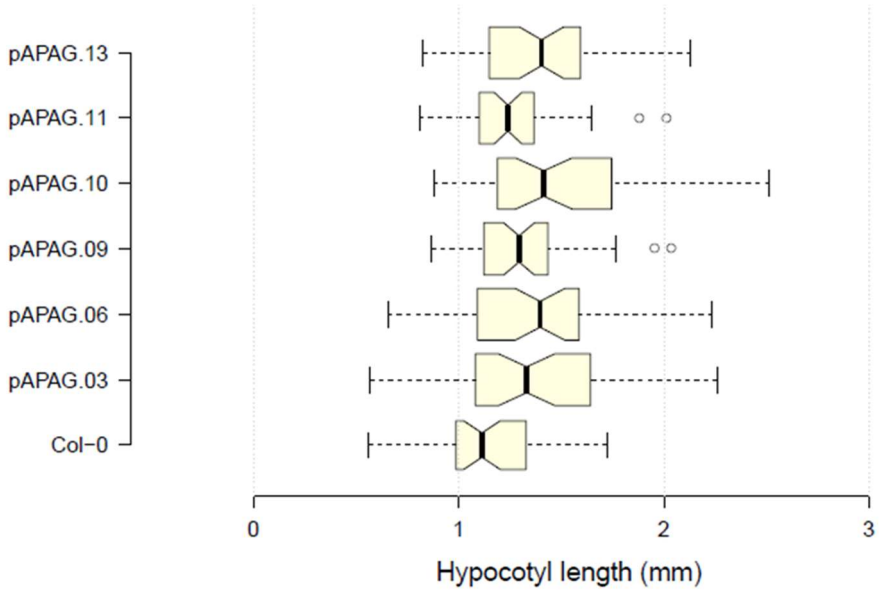

Figure 16 Boxplot diagram of the hypocotyl measurements of T_2 pAPAG plants under light (W) conditions. The notches are defined as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each

median. On the vertical axis it can be seen the different T_2 lines. On the horizontal axis, the length of the hypocotyl is shown in millimetres.

Figure 18 represents the results of the complementation test under simulated shade treatment. The complementation test had to be performed under simulated shade conditions because it is the only way to differentiate the phenotypes between the mutants and the wild-type (Figure 17).

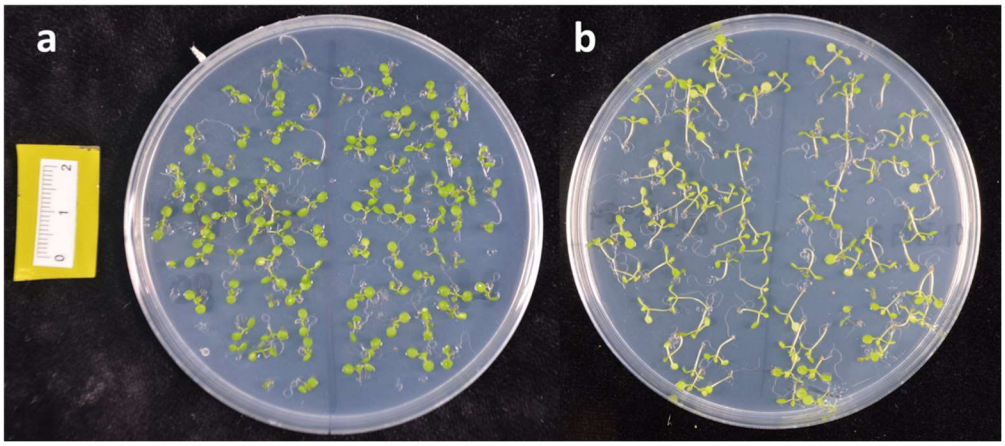

Figure 17. T_2 seedlings photographed in order to do hypocotyl measurements. In picture a, the seedlings were under W treatment. In picture b, the same seedlings but with simulated shade (W+FR) treatment.

Col-0 wild type under shade treatment, as Figure 18 shows, has a shorter hypocotyl than the *phyA* mutant. So, if complementation is achieved, the samples should behave as Col-0. The difference between Col-0 and *phyA* is highly noticeable. As the complementation test, the samples which seem to have a similar behaviour to Col-0 are pAPAG.14, pAPAG.13 and pAPAG.11. However, it cannot be seen very clearly. These three samples are the ones with more similar phenotype, but they are not fully mimicking Col-0 behaviour.

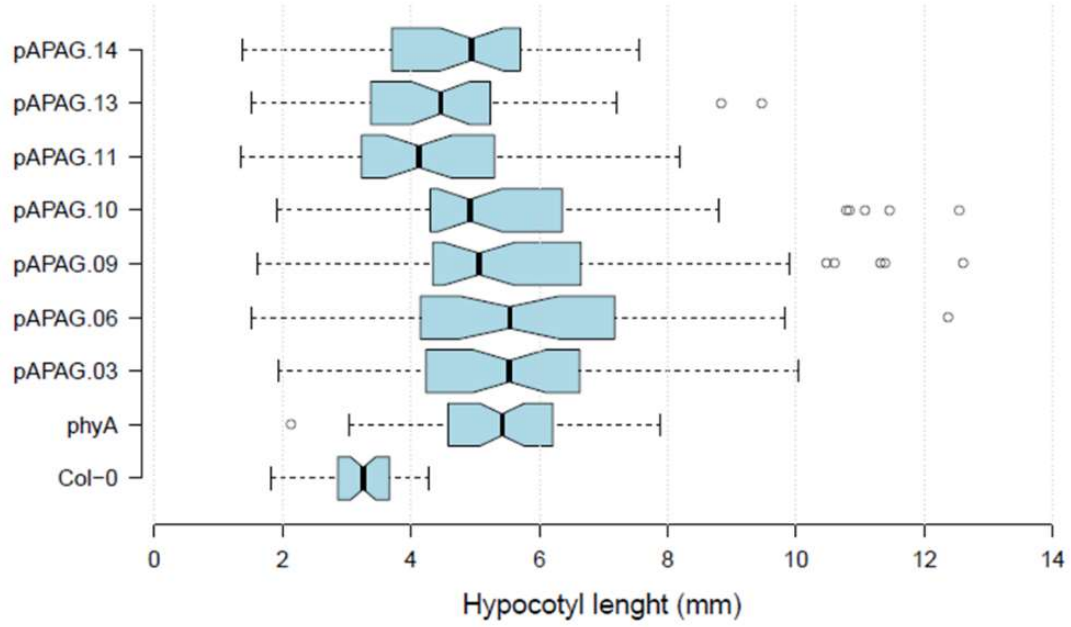

Figure 18 Boxplot diagram of the hypocotyl measurements of T_2 pAPAG plants under simulated shade treatment (W+FR). The notches are defined as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. On the vertical axis it can be seen the different T_2 lines. On the horizontal axis, the length of the hypocotyl is shown in millimetres.

In summary, for the 14 T_2 pAPAG lines analysed, the molecular analysis did not work. However, seven of these lines (pAPAG.03, pAPAG.06, pAPAG.09, pAPAG.10, pAPAG.11, pAPAG.13 and pAPAG.14) met the 3:1 segregation condition that indicate that they most probably contain 1 copy of the transgene. Moreover, three of these lines (pAPAG.11, pAPAG.13 and pAPAG.14) seem to partially complement the wild-type phenotype when analysed under simulated shade conditions. These three lines are good candidates to follow with the study.

4.4 Obtention of homozygous seeds for the pSP125 (*pHFR1:GFP-HFR1-3xHA*) construct in the *hfr1-5* mutant background

The construct that we used to analyse the expression of *HFR1* contains the promoter of *HFR1* (*pHFR1*) fused to the *GFP* reporter gene, the coding sequence of *HFR1* and a 3 histidine tag (AxHA). The 3xHA tag was added for other purposes different to the one that I studied. The name pSP125

is an internal short code for this construct that the host laboratory uses to simplify the nomenclature.

4.4.1 Selection for pSP125 T_2 sowing

As with the pAPAG T_2 samples, when I started working with this line with the T_2 generation of transformed plants was already available in the lab.

Also, as explained in section 3.2.3, in order for the T_2 samples to continue the process, its genetic segregation must follow a 3:1 proportion (75% resistant to PPT and 25% sensitive to PPT).

Table 6 Results of the selection of T_2 pSP125 plants. On the first column are all the 21 pSP125 T2 plants. In green are the ones which are considered transformed, due to the result of the Chi-square statistical test (χ^2 on the sixth column). In red are the ones dismissed from the experiment.

plants PPTR	not germinated	HgR	HgS	if $\chi 2<3.84$, 1 T-DNA	grow 8 PPT^R seedlings to pots
pSP125.01		83	18	1,71	YES
pSP125.02		76	33	1,62	YES
pSP125.03		54	35	9,74	NO
pSP125.04		103	0	34,33	NO
pSP125.05		83	20	1,71	YES
pSP125.06		73	24	0,003	YES
pSP125.07		67	34	4,04	NO
pSP125.08		68	29	1,24	YES
pSP125.09		94	8	16,01	NO
pSP125.10		83	20	1,71	YES
pSP125.11		89	9	13,07	NO
pSP125.12		65	30	2,19	YES
pSP125.13	DIED				
pSP125.14	DED				
pSP125.15		0	100	300	NO
pSP125.16		72	25	0,03	YES
pSP125.17		102	0	306	NO
pSP125.18		63	38	9,01	NO
pSP125.19		75	18	1,58	YES
pSP125.20		68	32	2,61	YES
pSP125.21		63	37	7,68	NO

21 T2 lines were analysed (Table 6). Unfortunately, lines pSP125.13 and pSP125.14 did not germinate. From the Chi-square test we conclude that the lines pSP125.01, .02, .05, .06, .08, .10,

.12, .16, .19 and .20 contain one single copy of the transgene and they are selected to proceed with the obtention of the T_3 generation. With the samples that can be used for the experiment, the molecular analysis was performed.

4.4.2 Genotyping of pSP125 T_2 plants

As already explained, genotyping could not be performed on T_1 plants. So, the molecular analysis is performed in pSP125 T_2 lines. Based on the results of the segregation analysis, the genotyping was done in samples pSP125.01, .02, .05, .06, .08, .10, .12, .16, .19 and .20. Figure 19 shows the first genotyping results obtained in these lines.

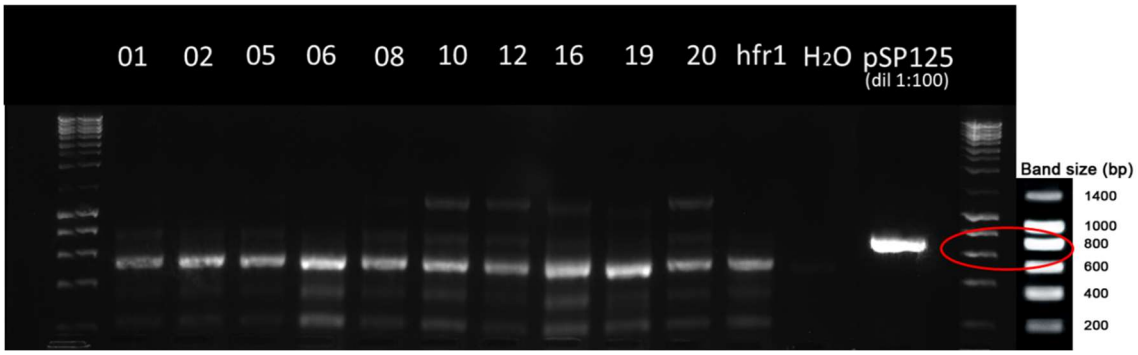

Figure 19 pSP125 PCR results of T_2 plants with SPO02 and SPO41 primers. 10 samples (pSP125.01, pSP125.02, pSP125.05, pSP125.06, pSP125.08, pSP125.10, pSP125.12, pSP125.16, pSP125.19 and pSP125.20) plus the positive and negative control (pSP125 diluted plasmid and H_2O respectively) and *hfr1* mutant.

For this PCR, the primers used were specific for the *HFR1* (SPO02) and for the GFP (SPO41)genes . Since the construct has the *HFR1* and the reporter *GFP* fused, the transformed samples should have incorporated both genes and therefore, it is possible to detect the transgene using these primers. The positive control showed the expected band (750 bp approximately). The negative control (H_2O) lane also worked since there is not any visible band. The 10 pSP125 T_2 lines showed all a detectable band, however the bands were not of the expected size. Moreover, a band of the same size was also present in the negative control *hfr1* mutant, indicating that those bands did not correspond with the presence of the transgene.

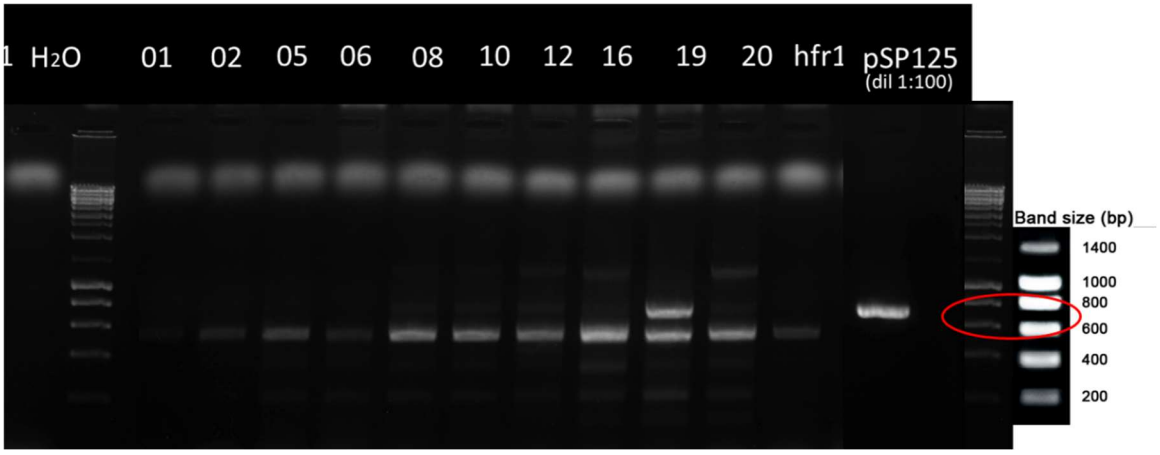

Figure 20 pSP125 PCR results of T_2 plants with SPO40 and SPO41 primers. 10 samples (pSP125.01, pSP125.02, pSP125.05, pSP125.06, pSP125.08, pSP125.10, pSP125.12, pSP125.16, pSP125.19 and pSP125.20) plus the positive and negative control (pSP125 diluted plasmid and H2O respectively) and *hfr1* mutant.

Since the first genotyping attempt did not work (Figure 19), a second PCR was performed with other primers. In this case, both primers were specific of the reporter GFP protein as in pAPAG T_2 genotyping (section 4.3.2). The results of the second PCR are presented in Figure 20.

In this case the positive control worked as well since there is a band with the expected size (approximately 750 pb). The negative control also worked since there is not a visible band on the gel. Still an unspecific band could be detected in all the lines, including the *hfr1* mutant. The only sample which coincides with the positive control is sample pSP125.19.

In conclusion, the presence of the transgene could only be confirmed by PCR in the pSP125.19 line.

4.4.3 Complementation test of the T_2 pSP125

As explained in section 4.3.3, the complementation test can be performed in T_2 seedlings since the resistant T_2 samples should have at least one copy of the transgene and in consequence, the phenotype of the mutant could be complemented.

Also, in order to have a comparison between the influence of the treatments, the complementation test has to be performed under W and W+FR conditions.

In Figure 21 are represented the results of the W complementation test.

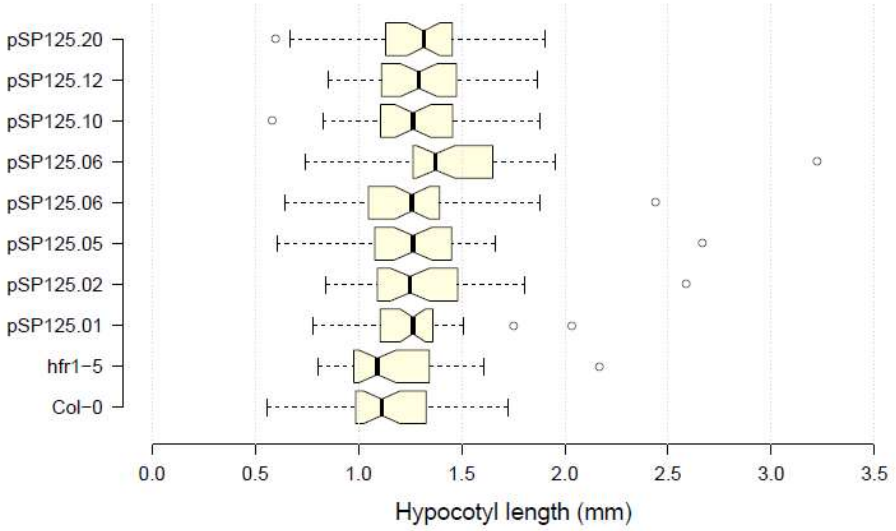

Figure 21 Boxplot diagram of the second attempt of hypocotyl measurements of pSP125 T_2 plants under light (W) treatment. The notches are defined as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. On the vertical axis it can be seen the different samples of T_2 plants. On the horizontal axis, the height of the hypocotyl is shown in millimetres.

In this experiment lines pSP125.16 and pSP125.19 were not included due to the lack of seeds.

Under this treatment, hypocotyl lengths seem to have a similar tendency. Here, the average in almost every sample is approximately between 1 mm to 1.5 mm. Also, there is not any measurement lower than 0.5 mm and greater than 2 mm excluding the outliers.

All the hypocotyl measurements under simulated shade of T_2 pSP125 lines are represented in Figure 22. In the boxes it can be found the 50% of the sample. Therefore, sample pSP125.08 has between 25% to 50% of the measures between approximately 4 mm to a little more than 5 mm. On the other hand, this same sample has between 50% to 75% of the measures between 5 mm to approximately 5.5 mm. Also, Col-0 seedlings display a shorter hypocotyl phenotype than *hfr1-5* under simulated shade conditions.

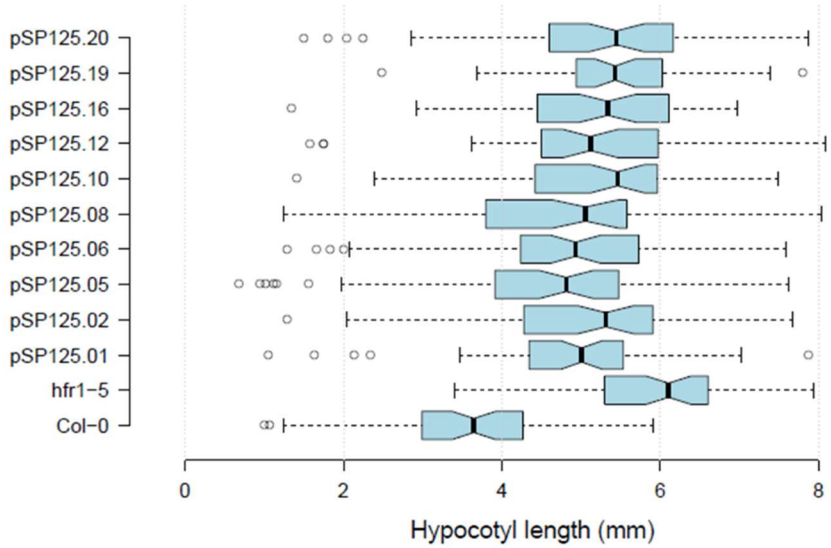

Figure 22 Boxplot diagram of the hypocotyl measurements of pSP125 T_2 plants under simulated shadow treatment (W+FR). The notches are defined as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. On the vertical axis it can be seen the different T_2 lines. On the horizontal axis, the length of the hypocotyl is shown in millimetres.

Overall, there is some complementation at first glance, as some measures of the transgenic lines overlap with the measures of the wild type, Col-0. Samples such as pSP125.08, pSP125.06, pSP125.05 and pSP125.02 seem to be more complemented than the others. Samples pSP125.08 and pSP125.02 have a wider range. In particular pSP125.08 which its hypocotyl length measures between 1.24 mm to 8.03 mm. Conversely, pSP125.01 or pSP125.19 are lines which are more alike to the *hfr1-5* mutant sample.

These experiments where actually done in parallel with the genotyping. Since the comparison between molecular results (Figure 19 and Figure 20) of the T_2 pSP125 were not conclusive, the analysis was repeated again. More T_2 seeds were sowed so that the PCR and the complementation test could be performed. As already mentioned, in this second experiment, there were not seeds left for lines pSP125.16 and pSP125.19, therefore they were not included in the analysis

The results of the second complementation test for T_2 pSP125 are represented in Figure 23.

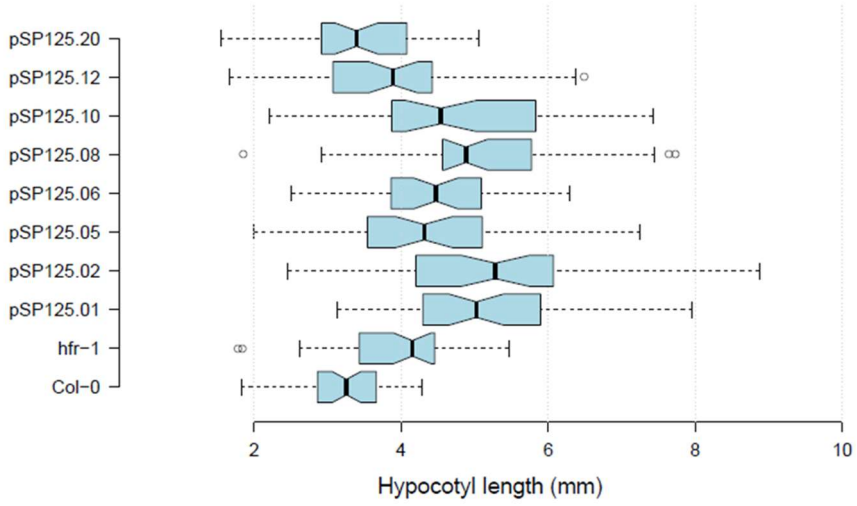

Figure 23 Boxplot diagram of the second attempt of hypocotyl measurements of pSP125 T_2 plants under simulated shadow treatment (W+FR). The notches are defined as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. On the vertical axis it can be seen the different samples of T_2 plants. On the horizontal axis, the height of the hypocotyl is shown in millimetres.

In this case, the results seem to be more shifted to the left, this is, the hypocotyls were in general shorter. As in Figure 22, lines pSP125.02, pSP125.05, pSP125.06 and pSP125.08 appears to be the better complemented, being samples pSP125.12 and pSP125.20 the ones more similar to Col-0.

In summary, for the 21 T_2 pSP125 lines analysed, 10 lines (pSP125.01, .02, .05, .06, .08, .10, .12, .16, .19 and .20) met the 3:1 proportion of segregation that indicate that they most probably contain at least 1 copy of the transgene (Figure 9). However, the molecular analysis could only confirm the presence of the transgene in the line pSP125.19. Moreover, from the complementation analysis, seems to not be complemented (Figure 22). Other lines, such as pSP125.02, .05, .06, and .08 they did show a partial complementation. These four lines are good candidates to continue with the study.

5 Conclusions

This bachelor thesis is made under an exceptional situation. Due to COVID-19 pandemic, it was not possible to go to the laboratory frequently and the working hours at CRAG were substantially reduced. Consequently, it was not feasible to fulfil all the objectives of the project. Despite of this, several advances were made in objectives 1 and 2, that I explain in the following conclusions:

- We have progressed in the selection and analysis of two of the three transgenic lines proposed in this work. Unfortunately, we did not transform *hy5* mutant plants with the pB71 (*pHY5:HY5-YFP*) construct, as we could not confirm the transformation of *A. tumefaciens* with the transgene.
- For the pAPAG plants (*pPHYA:PHYA-GFP* construct in the *phyA* mutant background), from the segregation analysis I identified 7 lines with a 1 copy of the transgene. However, the molecular analysis did not work.
- The complementation analysis in T_2 of pAPAG lines showed a partial complementation in 3 of the lines selected in the segregation analysis.
- For the pSP125 plants (*pHFR1:GFP-HFR1-3xHA* construct in the *hfr1-5* mutant background) , from the segregation analysis I identified 10 lines with a 1 copy of the transgene. However, the molecular analysis could only confirm the presence of the transgene in one of the lines (pSP125.19).
- The complementation analysis in T_2 of pSP125 showed a partial complementation in 4 of the lines selected in the segregation analysis.
- Segregation analysis might be better performed with homozygous plants (T_3 or T_4 generations).
- Since obtaining homozygous transgenic plants takes approximately 1 year it has not been possible to obtain them, neither to analyse the expression of the reporter genes.
- From a more personal point of view, it has helped me realise that science is sometimes not so obvious. And often it does not give us the results we expect. However, you learn that bad results are also results.

6 Problem solving and future perspectives

Since it was not possible to reach all of the three specific objectives, in this section I pretend to explain how I would continue the process.

For the pSP125 and pAPAG transgenic lines, I would start from the T_2 selection for sowing. I would start with a higher number of lines. I would perform the molecular analysis on T_2 seedlings, especially paying attention to the DNA extraction (trying to obtain DNA of higher quality) and the selection of primers. In summary, should optimize the genotyping.

I would perform the complementation tests with T_3 or T_4 homozygous plants. This should give me lower variability among samples, as well as more quantity of seeds to repeat experiments.

On the other hand, for the pB71 construct, I would develop the full process starting from the *A. tumefaciens* transformation. Performing a colony PCR from *A. tumefaciens* is challenging because of the cell wall of this bacteria, and sometimes the genotyping at this stage is not possible. To solve this, I could try to extract the plasmidic DNA from the bacteria to better detect it by PCR. Alternatively, I would anyway transform T_0 Arabidopsis with the *A. tumefaciens*, and then perform the molecular analysis on the T_1 seedlings. If the results are positive then, I would continue with the process. If not, I would have to transform Agrobacterium with pB71 construct again.

7 Bibliography

Andersen, T., Liang, D., Halkier, B., & White, R. (2014). Grafting Arabidopsis. BIO-PROTOCOL, 4(13). https://doi.org/10.21769/bioprotoc.1164

Ballaré, C. L., & Pierik, R. (2017). The shade-avoidance syndrome: Multiple signals and ecological consequences. Plant Cell and Environment, 40(11), 2530–2543. https://doi.org/10.1111/pce.12914

Boyes, D. C., Zayed, A. M., Ascenzi, R., Mccaskill, A. J., Hoffman, N. E., Davis, K. R., & Görlach, J. (2001). Growth Stage-Based Phenotypic Analysis of Arabidopsis: A Model for High Throughput Functional Genomics in Plants. In The Plant Cell, (13).

Casal, J. J. (2012). Shade Avoidance. The Arabidopsis Book, 10, e0157. https://doi.org/10.1199/tab.0157

Christie, J. M. (2007). Phototropin Blue-Light Receptors. https://doi.org/10.1146/annurev.arplant.58.032806.103951

Courbier, S., & Pierik, R. (2019). Canopy Light Quality Modulates Stress Responses in Plants. IScience, (22), 441–452. https://doi.org/10.1016/j.isci.2019.11.035

Edwards, K., Johnstone, C., & Thompson, C. (1991). A simple and rapid method for preparation of rhizopus oryzae genomic DNA for PCR analysis. Nucleic Acids Research, Oxford University Press, 19(6), 1349. https://doi.org/10.13031/aim.20141900217

Fairchild, C. D., Schumaker M. A., Quail P. H. (2000) HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction. Genes Dev, (14), 2377-2391.

Fisher, R. A. (1954). Statistical methods for research workers. p. 355. New York City.

Franklin, K. A., & Whitelam G. C. (2007). Light-quality regulation of freezing tolerance in Arabidopsis thaliana. Nat Genet (39), 1410–1413.

Galstyan, A., Cifuentes-Esquivel, N., Bou-Torrent, J., & Martinez-Garcia, J. F. (2011). The shade avoidance syndrome in Arabidopsis: a fundamental role for atypical basic helix-loop-helix

proteins as transcriptional cofactors. The Plant Journal, 66(2), 258–267. https://doi.org/10.1111/j.1365-313X.2011.04485.x

Koornneef, M., & Scheres, B. (2001). Arabidopsis thaliana as an Experimental Organism. In eLS. https://doi.org/10.1038/npg.els.0002031

Krämer, U. (2015). Planting molecular functions in an ecological context with arabidopsis thaliana. ELife, 2015(4). https://doi.org/10.7554/eLife.06100.001

Martínez-García, J. F., Gallemí, M., Molina-Contreras, M. J., Llorente, B., Bevilaqua, M. R. R., & Quail, P. H. (2014). The shade avoidance syndrome in Arabidopsis: The antagonistic role of phytochrome A and B differentiates vegetation proximity and canopy shade. PLoS ONE, 9(10). https://doi.org/10.1371/journal.pone.0109275

Molina Contreras, M. J. (2017). Regulación de las respuestas a la proximidad vegetal en Arabidopsis thaliana y especies afines. PhD thesis. (Universitat de Barcelona). Retrieved from http://www.tdx.cat/handle/10803/457872

Murasige, T., & Skoog, F. (1962). A Revised Medium for Rapid Growth and Bioassays with Tobacco Tissue Cultures. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Oyama, T., Shimura, Y., & Okada, K. (1997). The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus- induced development of root and hypocotyl. Genes and Development, 11(22), 2983–2995. https://doi.org/10.1101/gad.11.22.2983

Pedmale, U. V., Huang, S. S. C., Zander, M., Cole, B. J., Hetzel, J., Ljung, K., … Chory, J. (2016). Cryptochromes Interact Directly with PIFs to Control Plant Growth in Limiting Blue Light. Cell, 164(1–2), 233–245. https://doi.org/10.1016/j.cell.2015.12.018

Roig-Villanova, I. (2007). Control molecular i fisiològic de la síndrome de fugida de l'ombra en *Arabidopsis thaliana*. PhD thesis. (Universitat de Barcelona).

Roig-Villanova, I., & Martínez-García, J. F. (2016). Plant responses to vegetation proximity: A whole life avoiding shade. Frontiers in Plant Science. https://doi.org/10.3389/fpls.2016.00236

Runkle, E. (2016a). A Closer Look at Far-Red Radiation. *Greenhouse Product News*. Retrieved from https://www.canr.msu.edu/uploads/resources/pdfs/fr-radiation.pdf

Runkle, E. (2016b). Red Light and Plant Growth. *Greenhouse Product News*. Retrieved from https://www.canr.msu.edu/uploads/resources/pdfs/fr-radiation.pdf

Su, L., Hou, P., Song, M., Zheng, X., Guo, L., Xiao, Y., … Yang, J. (2015). Synergistic and antagonistic action of phytochrome (phyA and phyB) during seedling de-etiolation in arabidopsis thaliana. International Journal of Molecular Sciences, 16(6), 12199–12212. https://doi.org/10.3390/ijms160612199

Vervliet, G., Holsters, M., Teuchy, H., Van Montagu, M., & Schell, J. (1975). Characterization of different plaque forming and defective temperate phases in Agrobacterium strains. Journal of General Virology, 26(1), 33–48. https://doi.org/10.1099/0022-1317-26-1-33

Yang, S. W., Jang, I. C., Henriques, R., & Chua, N. H. (2009). Far-red elongated hypocotyl1 and FHY1-Like associate with the arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome signals for inhibition of hypocotyl elongation CW. Plant Cell, 21(5), 1341–1359. https://doi.org/10.1105/tpc.109.067215

Yu, Y., Wang, J., Zhang, Z., Quan, R., Zhang, H., Deng, X. W., … Huang, R. (2013). Ethylene Promotes Hypocotyl Growth and HY5 Degradation by Enhancing the Movement of COP1 to the Nucleus in the Light. PLoS Genetics, 9(12), e1004025. https://doi.org/10.1371/journal.pgen.1004025

Zhang, R., Yang, C., Jiang, Y., & Li, L. (2019). A PIF7-CONSTANS-Centered Molecular Regulatory Network Underlying Shade-Accelerated Flowering. Molecular Plant, 12(12), 1587–1597. https://doi.org/10.1016/j.molp.2019.09.007