Partnership throughout the ages:

The coevolution of the transcriptional regulators LEUNIG and SEUSS in the green plant lineage

Dissertation

for the degree Doctor of Science (Dr. rer. nat.),
submitted to the faculty for Biology and Chemistry,
Institute of Botany of the
Justus-Liebig-University Gießen

by

M. Sc. Julian Vincent Garrecht

1st assessor: Prof. Dr. Annette Becker
Institute of Botany
Justus-Liebig-University Giessen
Germany

2nd assessor: Prof. Dr. Günter Theißen
Department of Genetics
Friedrich-Schiller-University Jena
Germany

Examiner: Prof. Dr. Michael Frei
Agronomy and Crop Physiology
Justus-Liebig-University Giessen
Germany

Examiner: Prof. Dr. Sigurd Braun
Department of Genetics
Justus-Liebig-University Giessen
Germany

Zusammenfassung

Die Transkriptionsfaktor-Familien LEUNIG (LUG) und SEUSS (SEU) sind zentrale Modulatoren in der Entwicklung von Angiospermen, mit Funktionen in der Entstehung von Blüten, der männlichen und weiblichen Gametangien, der Embryonalentwicklung, sowie dem generellen Pflanzenwachstum. Ihre Interaktion formt eine regulatorische Plattform, welche Transkriptionsfaktoren und Histon-Modifikatoren zusammenbringt, um die Expression von Zielgenen sowohl zu aktivieren als auch zu reprimieren, und bilden dadurch ein weitreichendes genregulatorisches Netzwerk. Interessanterweise sind beide Familien stark konserviert und in allen Landpflanzenlinien sowie in vielen streptophyten Algen vertreten. Dies wirft Fragen über den Ursprung des LUG – SEU Moduls auf, sowie über seine Funktionen außerhalb der Angiospermen, und darüber, wie das Modul so evolvierte, dass es mehrere kritische, reproduktive Prozesse kontrolliert. Diese Dissertation zeigt, dass die Proteininteraktion zwischen den zwei Familien, und mit den MADS-Box Transkriptionsfaktoren, vor mindestens 800 Millionen Jahren in den streptophyten Algen evolviert ist, und dass LUG und SEU seither eine starke Koevolution aufweisen. Es wird demonstriert, dass sie als genregulatorische Schnittstelle fungieren, welche unter anderem zusammen mit den MADS-Box Transkriptionsfaktoren evolviert ist, um verschiedene Aspekte der Blütenentwicklung und Reproduktion in Angiospermen zu regulieren, gemeinsam mit weiteren Funktionen in der Stressregulation und den Signalwegen von Phytohormonen. Des Weiteren wurde die Funktion von LUG-Homologen im Blattmoos *Physcomitrium patens* untersucht, wodurch ihre Verbindung zu einem kritischen Entwicklungsübergang in Moosen, und zu der Regulation des Auxin-gesteuerten Signaltransduktionspfads in Landpflanzen offengelegt wurde.

Abstract

The LEUNIG (LUG) and SEUSS (SEU) families of transcriptional regulators are central modulators of angiosperm development, with relevance flower formation, development of male and female gametangia, embryo development, and general plant growth. Their interaction forms a regulatory platform, bringing transcription factors and histone modifiers together to both activate and repress target gene expression, thus creating a vast gene regulatory network. Interestingly, both families are deeply conserved, being present in all land plant lineages and in many streptophyte algae, raising questions about when the origin of the LUG – SEU protein module, about their functions outside of angiosperms, and how the module evolved to govern various critical reproductive processes. This dissertation demonstrates that the protein interaction between these two families, and with the MADS-box transcription factor family, evolved at least 800 million years ago in the streptophyte algae, an that LUG and SEU exhibit a strong coevolution ever since. They are shown to function as gene regulatory hubs that, among other things, evolved alongside the MADS-box proteins to regulate multiple aspects of flower development and angiosperm reproduction, in concert with functions in stress response and phytohormone signaling. Furthermore, the functions of LUG homologs in the moss *Physcomitrium patens* were studied, revealing their connections to a critical developmental transition in mosses, and to the regulation of the auxin signaling pathway of land plants.

Contents

1 Introduction

1.1 The evolution and terrestrialization of the green plant lineage

The terrestrialization of plants was arguably one of the most impactful events during plant evolution, not only opening up a vast range of new biological niches and drastically altering various aspects of earth itself, but also requiring extensive modifications in all areas of plant development, maintenance and reproduction, to cope with the new challenges and restrictions brought by the life on land. Among those are, for example, water scarcity, strong temperature changes both across the day and across seasons, or high solar and UV irradiance (Vries & Archibald, 2018; Kunz *et al.*, 2024). While multiple algae lineages independently evolved to live in terrestrial environments, dating as far back as at least 900 million years ago (MYA), only the embryophyta of the the green plants (viridiplantae) managed to establish themselves as the major plant macroorganisms on land (McCourt *et al.*, 2023; Vries & Archibald, 2018). Today, around 80 % of all biomass on earth belongs to plants, among which the embryophytes are by far the dominant group (Bar-On *et al.*, 2018).

The embryophyta are a division within the streptophyte algae, which additionally contains six major lineages of green algae (Fig. 1). Those lineages are sometimes grouped and referred to as the KCM-grade and the ZCC-grade, respectively, each containing three algae lineages (Bierenbroodspot *et al.*, 2024). Among them, the ZCC-grade Zygnematophyceae lineage represents the sister lineage to the embryophyta, whose most recent common ancestor (MRCA) lived around 600 MYA (Harris *et al.*, 2022). The emergence of the first embryophytic land plants has been dated to around 500 MYA (Harris *et al.*, 2022; Kapoor *et al.*, 2023; McCourt *et al.*, 2023), and is thought to be the result of gradual adaptation of fresh-water algae to land throughout the streptophyte lineage. Their ecotone habitat exerts a selective pressure, for example by cycles of drying and rehydration, that favors many traits crucial for survival on and adaptation to land (Kunz *et al.*, 2024; McCourt *et al.*, 2023), like more sturdy cell walls (Vries & Archibald, 2018), or improved protection against UV radiation (Kunz *et al.*, 2024). Additionally, elements of many critical regulatory and developmental networks found in land plants were already present throughout the streptophyte lineage, and modified to support life on land. For example, aspects of the auxin and cytokinin signaling pathways can be found in the KCM-grade Klebsormidiophyceae lineage (Hernández-García *et al.*, 2024; Hori *et al.*, 2014), and a high number of transcription factor (TF) families critical for land plant development were already present in extensive gene regulatory networks as well (Bowman, 2022; Vries & Archibald, 2018; Wilhelmsson *et al.*, 2017). Taken together, all these factors help to explain the unique position the embryophytes evolved into.

Early embryophyta therefore had many similar traits to their Zygnematophyceae sister lineage, like a multicellular haploid gametophyte life phase forming motile sperm and a diploid zygote, but key differences were also present at the incipient stage: Especially the

cycle between the gametophyte producing male antheridia and female archegonia that give rise to an embryo, which develops into a multicellular diploid sporophyte generation that in turn produces haploid spores, was likely a crucial advantage in sexual reproduction in comparison to algae (Becker *et al.*, 2025; Bowman, 2022). Similarly, the evolution of rooting structures, or of apical meristems that allow three dimensional growth were further differences aiding the terrestrialization of early embryophytes (Bowman, 2022; Moody, 2022). These novelties required the extensive modification and expansion of present gene regulatory networks, to allow for the emergence and regulation of new genes and expression patterns, resulting in a variety of novel vegetative and reproductive organs throughout land plant evolution. Especially whole genome duplications (WGDs) are considered prime drivers of evolutionary novelties and of speciation, as they allow for the large-scale expansion of TF families and subsequent sub- and / or neofunctionalization of genes within these regulatory networks (Kapoor *et al.*, 2023; van de Peer *et al.*, 2009).

During land plant evolution, the embryophyta lineage split into two monophyletic clades, namely the bryophyte lineage containing the hornworts, liverworts, and mosses, and the tracheophyte lineage containing the lycophytes, monilophytes, and the seed plants, which consist of the gymnosperm and angiosperm lineages (Fig. 1). Many bryophyte species still exhibit early terrestrial adaptations, such as a strongly water-dependent fertilization or a gametophyte-dominant life phase where the sporophyte depends on the maternal gametophytic tissue for survival (Becker *et al.*, 2025; Rensing *et al.*, 2020). During the evolution of the tracheophytes, also called vascular plants, on the other hand, various organs, mechanisms and structures emerged that reduced these dependencies and constraints, and increased the viability and variability of land plants even more: In lycophytes and monilophytes, both the gametophyte and the sporophyte phases are free-living and largely independent from each other, besides during the first developmental stages of the sporophyte. Additionally, their sporophytes usually show a higher degree of complexity and develop much larger than the respective gametophytes. In seed plants, this disparity increased even further, with the sporophyte becoming the main life phase of the plant that nourishes a strongly reduced gametophyte and protects it in specialized structures like pollen or ovules, which can double as dispersion mechanisms. This is especially prominent in the flowers of angiosperms, which added an additionally layer of protection in the form of carpels to the ovules and later the seeds, but also increased reproductive success by capturing large quantities of pollen, allowing parallel fertilization, and enhancing seed dispersal (Becker *et al.*, 2025; Bowman, 2022).

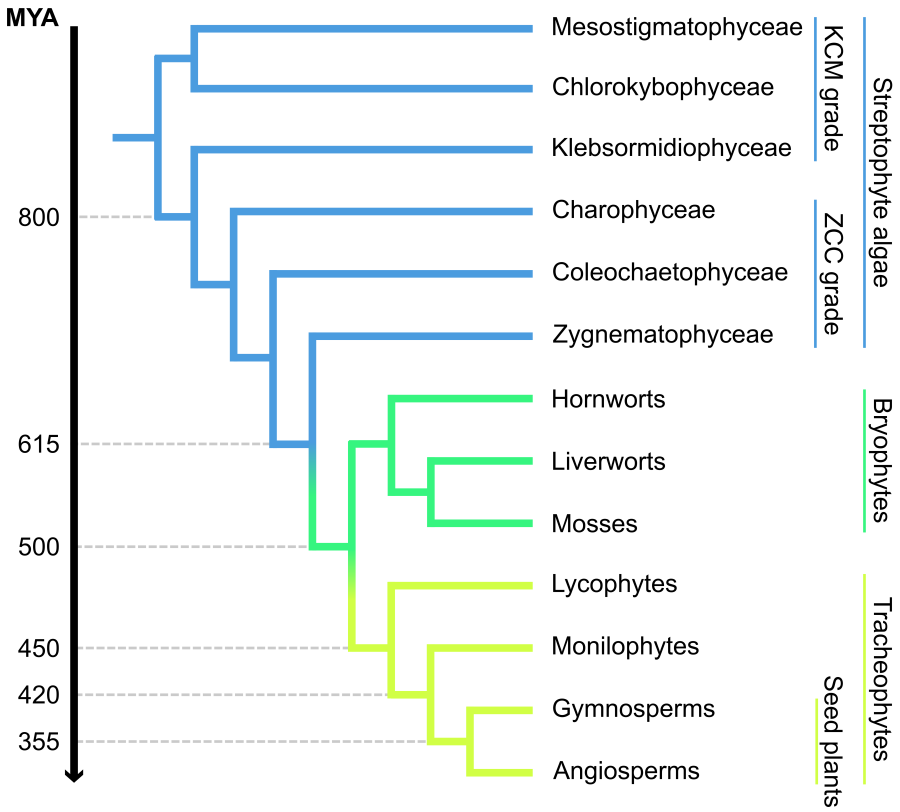


Figure 1: Phylogeny of major streptophyte and embryophyte lineages. The major lineages of streptophyte algae are colored blue, while bryophyte lineages are colored green, and tracheophyte lineages are lime (Bierenbroodspot *et al.*, 2024; G.-Q. Liu *et al.*, 2022). Estimated crown ages for selected branches based on Harris *et al.* (2022).

1.2 Representative species of major green plant lineages

To better understand the evolution of gene families, regulatory networks, and developmental pathways, as well as the emergence of novel organs and strategies for survival underlying the evolution of the green plant lineage and especially the embryophyta, much effort was put into researching and establishing model organisms for major streptophyte lineages. Especially the creation of quality reference genomes and transcriptomes for a multitude of species proved to be invaluable for the comprehension of (land) plant evolution, by allowing tracing, analyzing, and understanding the evolutionary and developmental trajectories of specific plant organs, cellular structures, or of specific gene families. In the following section, model species for several major plant lineages that were used in this dissertation will be briefly presented, to gain an understanding about the structural and genetic features and peculiarities of both the specific species, and the different lineages they represent.

1.2.1 *Klebsormidium nitens*, a model streptophyte algae

The streptophyte algae *Klebsormidium nitens*, formerly *Klebsormidium flaccidum*, was the first streptophyte algae with a fully sequenced reference genome (Bierenbroodspot *et al.*, 2024; Hori *et al.*, 2014). It belongs to the order Klebsormidiales, which resides within

the Klebsormidiophyceae of the paraphyletic KCM-grade of streptophyte algae, and diverged from the lineage leading to land plants around 800 MYA (Fig. 1). *K. nitens* forms multicellular, non-branching filaments without specialized cell types and a singular, large chloroplast; and like other Klebsormidiophyceae, lives in both freshwater and terrestrial habitats (Hori *et al.*, 2014).

Its genome is around 120 megabases (Mb) in size, with ~ 16.000 protein coding genes currently known. Many gene families or functional protein domains critical for land plants can already be found in *K. nitens*, such as factors connected to multicellularity, phytohormone signaling, or stress response mechanisms, and especially those relating to drought or high light stress (Hori *et al.*, 2014; Serrano-Pérez *et al.*, 2022; Wilhelmsson *et al.*, 2017).

1.2.2 Development of and hormonal signaling in the moss *Physcomitrium patens*

Bryophytes are the sister lineage to vascular plants, and as such in a great position to study both the transition from water to land, and the evolution of many tracheophyte characteristics (Fig. 1). The moss *Physcomitrium patens*, formerly *Physcomitrella patens*, has been used to this effect for over 50 years, due to its simplistic body plan, small size, and fast propagation, as well as the general availability of a multitude of different genetic tools and resources. It belongs to the family of the Funariaceae and emerged presumably 10 MYA, living mainly in the northern hemisphere (Rensing *et al.*, 2020). Its genome has an estimated size of 518 Mb, distributed across 26 or 27 chromosomes, and with ~ 30.000 protein coding genes. Around 60 % of the genome consists of repeats; and two ancestral WGDs that both happened during the last 70 million years could be identified (Bi *et al.*, 2024; Lang *et al.*, 2018).

1.2.2.1 The life cycle of *P. patens*

As a bryophyte, the dominant life cycle phase of *P. patens* is the gametophyte, which develops from a haploid spore. Upon germination, a one dimensional, tip-growing protonema cell emerges, that quickly forms a two dimensional (2D) filamentous outgrowth (Moody, 2022; Rensing *et al.*, 2020). Multiple types of protonema cells exist, among which chloronema and caulonema comprise the two principal forms: Chloronema are characterized by being primarily photosynthetically active cells, containing a high amount of chloroplast, and exhibiting a transverse cell plate between chloronema cells. Caulonema on the other hand are colonizing cells that show fast outgrowth, are thinner with only few chloroplasts, and possess oblique cell plates (Jaeger & Moody, 2021; Rensing *et al.*, 2020). The first emerging cells of germinating spores are in fact chloronema that grow from an apical chloronema stem cell. This tip growing stem cell eventually transitions into a caulonema apical stem cell that enables 2D growth by forming side branch initial cells along the filament that either become chloronema or caulonema apical stem cells, or, in around 5 % of cases, form gametophore producing buds (Moody, 2022; Rensing *et*

al., 2020). Additionally, chloronema cells not adjacent to apical stem cells can also form side branches and become secondary apical stem cells, but are not able to initiate gametophore growth (Kofuji & Hasebe, 2014).

The three dimensional (3D) growth of gametophores progresses in three stages: Firstly, a gametophore apical cell is initiated from a caulonema side branch, which is signaled by a change of the cell division plane, a process that is highly localized, but unfortunately still poorly understood. Secondly, the gametophore initial develops into a tetrahedral cell capable of producing multiple daughter cells downwards, thus enabling 3D growth. Finally, the apical stem cell begins cell division and produces merophytes, rising above the protonema in the process (Moody, 2022). The resulting gametophore consists of a gametophore stem with ovate, lanceolate leaflets in a spiral phyllotaxis that are comprised of a single cell layer, except for their median region, as well as specialized cells for water and nutrient transport (Rensing *et al.*, 2020). At the base of the gametophores, rhizoids emerge, which are a protonema variant specialized for rooting; they share many similarities with caulonema, but contain immature plastids and show a brown pigmentation (Kofuji & Hasebe, 2014).

Under short day conditions, the formation of gametangia at the gametophore apex is induced. As *P. patens* is a bisexual species, both antheridia and archegonia develop at the apices, with the sperm cell producing antheridia emerging first, followed by the appearance of archegonia, with the archegonial neck rising above the antheridia. The fertilization itself is water dependent, as it requires the sperm cells to swim towards and through the archegonial neck. Afterwards, the fertilized, diploid embryo develops into the spore producing sporophyte, which is anchored to and depends on the gametophore for nutrient transfer. The immature sporophytes appear green, while adult sporophytes are brown and can release mature, haploid spores (Rensing *et al.*, 2020).

1.2.2.2 Developmental transitions in *P. patens* are tightly regulated by the phytohormones auxin and cytokinin

Both the transition from chloronema to caulonema identity, and the initiation of gametophore apical stem cells is tightly regulated by a plethora of genes and phytohormones: The auxin and cytokinin signaling pathways function as general regulators of these transitions, while nutrient availability in the medium and the photosynthetic potential, mediated via phytochrome and cryptochrome pathways, are influential factors as well (Jaeger & Moody, 2021; Kofuji & Hasebe, 2014).

The phytohormone auxin is consistently linked to cell division, expansion, and differentiation throughout land plants. During the chloronema – caulonema transition of the apical filament cells, localized auxin maxima are generated by both auxin synthesis and through polar auxin transport via PIN-FORMED (PIN) efflux carriers (Jaeger & Moody, 2021; see section 1.6). These auxin maxima activate the expression of two sets of genes,

namely of the *ROOT HAIR DEFECTIVE SIX-LIKE 1* (*PpRSL1*) and *PpRSL2* genes, and of the *LOTUS JAPONICUS ROOTHAIRLESS1-LIKE 1* (*PpLRL1*) and *PpLRL2* genes. All four genes transcribe for basic helix-loop-helix (bHLH) TFs that are critical for establishing caulonema identity along the protonema filament. Interestingly however, while their combined activity is required to correctly promote caulonema identity, both families act independently from each other (Jaeger & Moody, 2021; Jang & Dolan, 2011; Tam *et al.*, 2015; Thelander *et al.*, 2018).

Cytokinin on the other hand is often found as an antagonist to auxin, and indeed, the chloronema – caulonema transition is suppressed by the activity of cytokinin (Ashton *et al.*, 1979). Instead, cytokinin promotes secondary branching of chloronema and the 3D growth transition from caulonema to gametophore buds, two processes that are in turn inhibited by auxin (Moody, 2022; Thelander *et al.*, 2005). Notably however, formation of the gametophore initial can't be solely triggered by cytokinin, as low levels of auxin are required as well, to correctly orient the cell division plane (Moody, 2022).

The exogenous application of auxin or cytokinin results in dosage-dependent transitional defects: Elevated auxin conditions inhibit gametophyte growth by severely reducing or even abolishing the formation of chloronema, instead resulting in enhanced caulonema formation, which are often strongly pigmented. Gametophore production is decreased as well, with most developing gametophores being malformed. At high auxin concentrations, gametophore induction is completely abolished (Ashton *et al.*, 1979). Conversely, exogenous cytokinin suppresses caulonema differentiation and increases secondary chloronema branching, except under very high concentrations, where branching is inhibited instead. The production of gametophores is notably increased, but depending on the cytokinin dosage, strong deformations are present, resulting in callus-like buds of tissue that fail to develop shoots or rhizoids, but sometimes exhibit the outgrowth of leaflets (Ashton *et al.*, 1979; Thelander *et al.*, 2005).

Additionally, while the exact influence of auxin and cytokinin on gametangia development is still poorly understood, it has been shown that the regulators of auxin biosynthesis SHORT INTERNODE (SHI)/STYLISH (STY) are critical for correct antheridia and archegonia development (Landberg *et al.*, 2013; Thelander *et al.*, 2018). Further, many more cellular processes and genes families besides the ones described above are implicated in the developmental transitions of *P. patens*: For example, actin and actin-related factors are critical for correct cytoskeleton organization, the *NO GAMETOPHORES* regulators and the APB family of TFs play a crucial role for the formation of the gametophore apical cell (Moody, 2022), and the gibberellin signaling pathway has been shown to act in concert with auxin in the chloronema to caulonema transition (Jaeger & Moody, 2021). Thus, these developmental transitions are tightly controlled to ensure correct timing and proper functionality under favorable conditions.

1.2.3 *Marchantia polymorpha*, a representative liverwort species

Marchantia polymorpha is a cosmopolitan model liverwort species of the class Marchantiopsida, which has been studied extensively for at least 200 years. It is a dioecious species with separate male and female plants and a short life cycle exhibiting a dominant gametophytic phase: Starting from a germinating haploid spore, a filamentous protonema forms that develops into a thalloid main body, which in turn expands from notches at the thallus periphery. Propagation can either occur asexually via gemmae produced dorsally on the thallus in specialized structures called gemmae cups, or sexually via the production of male or female gametangia. The antheridia or archegonia are stalked and lifted above the thallus on antheridiophores or archegoniophores, respectively, and depend on water to both release motile sperm cells and to allow for fertilization of the egg cell. Within the archegonium, the fertilized zygote develops into a sporophyte that is dependent on the archegoniophore for nutrient transfer. Following meiotic divisions, male and female spores are discharged from the sporophyte (Sauret-Güeto *et al.*, 2020; Shimamura, 2016).

The genome of *M. polymorpha* consists of around 225 Mb across eight autosomes and one of two sex chromosomes. It encodes for ~ 19.100 protein coding genes, with approximately 5.400 alternative transcripts and circa 22 % repetitive elements. Interestingly, no WGDs in the lineage leading to *M. polymorpha* were recorded (Bowman *et al.*, 2017).

1.2.4 The novel genetic model fern *Ceratopteris richardii*

The fern *Ceratopteris richardii*, also known as C-fern, is a tropical homosporous leptosporangiate fern of the family Pteridaceae in the lineage of the monilophytes, the closest sister lineage to seed plants (Fig. 1). Both its haploid and diploid generations are living independently from each other, but show large differences in size and structural complexity. Haploid *C. richardii* spores develop into either hermaphroditic gametophytes carrying both antheridia and archegonia, or into smaller male gametophytes only exhibiting antheridia. The release of motile sperm and the fertilization of the egg cells in the archegonium are both water-dependent. The resulting sporophytes are dependent on the gametophyte only during their first few developmental stages, after which they develop into considerably larger, free-living plants. At first, sporophytes start to develop small, lobed, sterile fronds, which become successively larger and more dissected, culminating in the development of large fertile fronds that bear the spore generating sori (Conway & Di Stilio, 2020). The whole life cycle can be completed in approximately four months, and together with the relatively easy growth in lab or greenhouse environments, the high amount of descendants, and possibilities for self-fertilization, underscore the importance of *C. richardii* a model fern species since the 1970s (Hickok, 1977).

The genome of C-fern is 9.6 gigabases (Gb) large and organized in 39 chromosomes. Approximately 85 % of the genome are repetitive elements, and after two WGDs during the last 300 million years, it contains around 36.800 protein coding genes with a notably

high variation in intron lengths. Interestingly, several homologs of genes critical for angiosperm flower and seed development can be traced in C-fern, highlighting the conservation of these gene families and their recruitment into novel regulatory processes throughout evolution (D. B. Marchant *et al.*, 2022).

1.2.5 *Arabidopsis thaliana* as a model for flower development

Arabidopsis thaliana is a herbaceous, dicotyledon angiosperm of the family brassicaceae, and a widely used seed plant model. It is an annual plant of small stature, with a fast generational cycle and predominantly found in western Eurasia (Krämer, 2015). Its fully sequenced genome consists of 125 Mb DNA organized in five chromosomes, with around 26.200 protein coding genes, and multiple detectable ancestral whole genome dupli- and triplication events (Bowers *et al.*, 2003; The Arabidopsis Genome Initiative, 2000).

As a seed plant, its dominant life cycle phase is the sporophyte: Starting from a germinating seed, both a primary root and the cotyledons atop the hypocotyl emerge, which firstly develop a vegetatively growing rosette of leaves that later gives rise to an inflorescence-bearing shoot. Each inflorescence produces several terminal, bisexual flowers consisting of four whorls, centripetally called sepals, petals, stamens, and carpels. The two inner whorls develop the male and female gametophytes, respectively, which are strongly reduced and only consist of a few cells each: Stamen are comprised of anthers producing pollen grains and are stalked by the filaments, while the carpels are fused to the gynoecium. In *A. thaliana*, it is divided into two valves that are separated by a false septum containing placental tissue, from which ovules develop. These ovules bear the female gametophyte, which is protected by two integuments, and entirely dependent on the surrounding sporophyte for nutrient transfer. Fertilization is mediated by two distinct structures atop the gynoecium, the stigmatic papillae and the style: After pollination, the pollen germinates at the stigma, forming a pollen tube that grows through the transmitting tract in the style towards the ovules, where fertilization occurs. The gynoecium then develops into a silique that is connected to the plant via the pedicel, while the ovules develop into seeds. After seed maturation, the silique ruptures to release the seeds (Alvarez-Buylla *et al.*, 2010; Becker *et al.*, 2025; Crawford & Yanofsky, 2008).

1.3 Protein network dynamics and the mechanisms of coevolution

The evolution of plants and of life in general is largely driven by the kinetics of and interplay between genes, proteins, and other regulatory elements. Understanding their evolutionary dynamics and trajectories is a useful way to trace not only the evolution of singular genes or larger interaction networks, but can also be used to gain knowledge about the origin, development, and regulation of larger structures like specific tissues, organs, and even whole species. Especially the study of protein interaction networks is an effective avenue to deepen our perception of such network dynamics, and of coevolution in

the context of protein families.

Protein interactions can be generally characterized as networks consisting of nodes (proteins) that are connected by edges (interactions). They follow a „scale-free“ topology, where a high number of proteins is only interacting with a few others, while a few hubs proteins contain a significant amount of the total interactions in the network. This topology is also called a power law distribution (Albert & Barabási, 2002; Barabási, 2013; Ng & Huang, 2004; Ramos *et al.*, 2024; Stelzl *et al.*, 2005), and is based on continuous growth of the network and thought to be achieved by preferential attachment of new nodes to those already well connected (Albert & Barabási, 2002; Barabási, 2013; Eisenberg & Levanon, 2003). Alternatively, a „crystal growth“ model has been proposed, where the connection probability correlates with the availability of unoccupied surfaces in an interaction cluster, which takes the age of proteins into consideration better than other models (W. K. Kim & Marcotte, 2008). Regardless of the exact underlying mechanism, protein interaction networks follow several general behaviours: (1) older parts of the networks are better and denser connected than younger parts, and as such, newer proteins tend to preferentially interact with these hubs, thus further increasing their size (Batada *et al.*, 2006; Eisenberg & Levanon, 2003). (2) Protein networks are more resistant against arbitrary loss of nodes than a random network would be, but are highly impacted by the loss of hub proteins (Albert & Barabási, 2002; Ramos *et al.*, 2024). (3) Networks are „small worlds“, and as such, both the distance between any two nodes and the diameter of the whole network are relatively small, with the average path length between two proteins being 3.5 to 4 across multiple different databases (Albert & Barabási, 2002; Ramos *et al.*, 2024).

Rewiring of protein networks is a critical driver of evolution, and can happen in two different ways: either by gaining and loosing whole proteins, or only specific interactions between them. These changes can for example happen by gene duplication, which is often followed by a rapid divergence of one of the doubled genes, and therefore of the inherited interactions. Interestingly, in many cases, a portion of the original interactions are retained during the divergence (Andreani & Guerois, 2014; Sun & Kim, 2011). Alternatively, mutations within proteins can alter their structure, function, or expression, and thus lead to loss or gain of specific interactions. While the vast majority of possible mutations are prohibited at any given time due to deleterious effects on protein fitness, they can be compensated by either permissive neutral mutations allowing for more flexibility at other sites, or by mutations stabilizing negative effects (Andreani & Guerois, 2014; Lovell & Robertson, 2010; Povolotskaya & Kondrashov, 2010). Such changes can happen both within one sequence, as well as between interacting proteins, and are defined as intra- or intermolecular coevolution, respectively.

Generally, coevolution is often described as „reciprocal evolutionary changes in interacting species“ or, more specifically in the context of protein interactions, as „reciprocal evo-

lutionary changes in evolutionary interacting loci“ (Lovell & Robertson, 2010; Thompson, 1994). Thus, changes in the interaction interface of one protein of an interacting pair can be compensated by mutations of the partner protein, causing both proteins to coevolve due to their shared interaction. Alternatively, coevolution can also be visible in strongly conserved molecular sites, when one location exerts strong selective pressure on another, causing both sites to barely change across vast evolutionary distances (Lovell & Robertson, 2010). Indeed, especially interaction interfaces were observed to not only coevolve with each other, but also to be more conserved than other protein regions (Andreani & Guerois, 2014; Lovell & Robertson, 2010; Rodriguez-Rivas *et al.*, 2016). Conversely, intrinsically disordered regions (IDRs) are regions of proteins lacking fixed interfaces, instead relying on structural variability and malleability to interact with a large number of different protein partners. They are adept at enabling, coordinating, and modifying a plethora of interactions not only with other proteins, but also with DNA, RNA, or signaling molecules, and can affect the general behaviour of their proteins in a variety of ways. For example, they often exhibit small linear motifs that convey different functions, such as post-translational modifications or binding affinity for specific domains or proteins, whose specificity or strength can be further modified by a range of other aspects (Holehouse & Kragelund, 2024; Már *et al.*, 2023; van der Lee *et al.*, 2014). As such, IDRs exhibit a high sequence divergence, and interactions conveyed by them are prone to fast rewiring and modification (Andreani & Guerois, 2014; Már *et al.*, 2023).

1.4 The transcriptional regulators LEUNIG and SEUSS

One common denominator of life at large is the transfer of information between two entities, which can happen in a multitude of ways and with various kinds of information. Especially the vertical transfer of genetic information via reproduction is a prominent expression of this principle, and often tightly regulated to ensure the correct inheritance of such information (Skern-Mauritzen & Mikkelsen, 2021). In plants, reproduction can happen asexually and sexually, with the latter being a critical mechanism for the introduction of genetic exchange between individuals of a population, which has the potential to increase the viability and overall fitness of the population (Wagner, 2010).

The flowers of angiosperms play a crucial role in reproduction, and as such need to be tightly controlled in both development and functionality, to ensure the proper formation of the filial generation. Among the vast number of genes, protein, and other factors relevant for flower formation and reproductive success, the transcriptional regulators LEUNIG (LUG) and SEUSS (SEU) are known to be critical regulators of many of them, with their activity ensuring, among other things, the correct organ formation of the flower, of both male and female gametangia, and of the embryo (Azhakanandam *et al.*, 2008; Shi *et al.*, 2024; Sridhar *et al.*, 2004; Stahle *et al.*, 2009).

1.4.1 Structural characteristics of LUG and SEU proteins

The LUG family in *A. thaliana* consist of LUG and LEUNIG-HOMOLOG (LUH), two glutamine (Q)-rich proteins that are characterized by a „LUG/LUH, Flo8, Single-Strand DNA-binding Protein (SSDP)“ (LUFS) domain at their N-terminal, and a propeller-like structure of seven WD40 repeats at their C-terminal end, which are separated by an intermediate region (Fig. 2; Conner & Liu, 2000). The LUFS domain shares similarities with domains of the yeast Flo8 protein and human SSDP, consisting of a „lissencephaly homology“ (LisH) dimerization motif that can be found across eukaryotes, followed by a conserved sequence motif (Z. Liu & Karmarkar, 2008; Sridhar *et al.*, 2004). It is critical and sufficient for dimerization with the common interaction partner SEU, but also allows for formation of homotetramers (Sridhar *et al.*, 2004; H. Wang *et al.*, 2019). WD40 domains on the other hand are a class of highly abundant domains facilitating the formation of larger protein complex with various different interaction partners (Xu & Min, 2011). Interestingly, some protein interactions between LUG and other partners also require the LUG intermediate region to function correctly (Gonzalez *et al.*, 2007; Vega-León *et al.*, 2025).

Due to their similar domain structure and analogous functions, the LUG family is counted to the „Groucho (Gro)/Transducin-like Enhancer (TLE)“ super family of transcriptional corepressors, with high similarity to the Gro protein in fruit flies, TLE in mammals, or TUP1 in yeast. However, these similarities do not extend to their protein sequences (Conner & Liu, 2000; Z. Liu & Karmarkar, 2008). Members of this super family are global gene regulators lacking DNA binding domains, instead requiring interaction with cofactors for targeted gene regulation in mainly developmental and stress response pathways (J. E. Lee & Golz, 2012). In plants, a second group of Gro/TLE proteins are present, namely the TOPLESS (TPL) and TOPLESS-RELATED (TPR) family, which also possess a LisH domain, but lack the conserved sequence motif of the LUFS domain (J. E. Lee & Golz, 2012; Z. Liu & Karmarkar, 2008).

SEU and the three SEUSS-LIKE (SLK) proteins in *A. thaliana* are Q-rich proteins consisting of a centrally located „LIM domain-binding“ (LDB) domain that is flanked by two IDRs in the N- and C-termini (Fig. 2; Franks *et al.*, 2002; B. Wang *et al.*, 2022). As such, the SEU family is structurally similar to LDB proteins of animals and yeast, a group of transcriptional adaptors lacking DNA binding sites, thus require DNA-binding cofactors for target gene regulation. While members of this group can interact with both LIM and LUFS domain containing proteins, SEU and the SLKs lack an additional LIM interaction domain found in other LDB proteins and are as such unable to interact with plant LIM domain proteins (Bao *et al.*, 2010; Franks *et al.*, 2002; Sridhar *et al.*, 2004; van Meyel *et al.*, 2003). Therefore, the SEU family primarily interacts with the LUG family, connecting specificity-conveying TFs to transcriptional regulators (Bao *et al.*, 2010; Sridhar *et al.*, 2004).

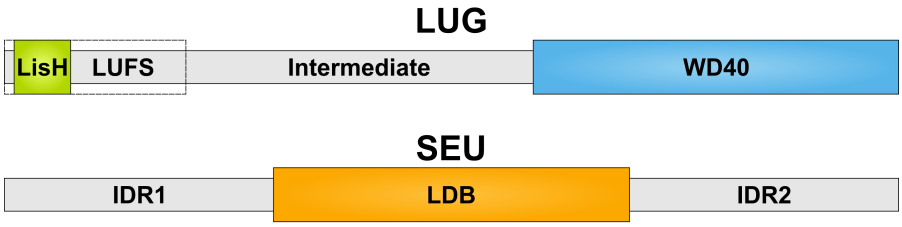


Figure 2: Protein domain organization of LUG and SEU proteins. LUG proteins (top) consist of the N-terminal LUFS domain that is sub-divided into the LisH motif (lime) and a conserved sequence (black rectangle). The WD40 domain (blue) is C-terminally located. SEU proteins (bottom) exhibit a centrally located LDB domain (orange) flanked by two IDRs.

1.4.2 The mechanisms of transcriptional regulation by LUG and SEU

LUG and SEU proteins can both activate and repress target gene transcription, either by acting independently from each other, or by forming a regulatory complex, the LUG – SEU module. In both cases, the specificity for and exact regulation of particular target genes is mediated by their various binding partners, while LUG and SEU act as a gene regulatory platform, bringing the different components together (N. Lee *et al.*, 2015; Sridhar *et al.*, 2004; Zhai *et al.*, 2020).

1.4.2.1 Formation of the LUG – SEU module

The interaction between LUG and SEU proteins is facilitated by binding of the SEU protein to the LUFS domain of LUG. While the LUFS domain alone is sufficient for successful binding, the whole SEU protein is required, and is only able to tolerate a small loss of its C-terminal IDR2 (Sridhar *et al.*, 2004). This general interaction mechanism is deeply conserved: In animals, SSDP can interact via their LUFS domain with the LDB containing LDB1 protein to form higher order protein complexes and regulate target gene expression; however, this interaction does not require fully intact LDB proteins (Bao *et al.*, 2010; van Meyel *et al.*, 2003).

In *A. thaliana*, interactions between LUG and SEU proteins have been shown for all homolog combinations, as both LUG and LUH are able to interact with all four members of the SEU family (Shrestha *et al.*, 2014; Stahle *et al.*, 2009). *In planta*, members of both families have been shown to be noticeably enriched in immunoprecipitation and mass spectrometry experiments using LUH as target protein, as well as being parts of a core hub in early floral developmental protein interaction networks (Vega-León *et al.*, 2025).

1.4.2.2 Canonical target gene regulation by the LUG – SEU module

The combined LUG – SEU module mainly functions as a transcriptional repressor of target genes during plant development: In *A. thaliana*, TFs like the MADS-box proteins APETALA 1 (AP1) or SEPALLATA 3 (SEP3) bind to SEU and convey target specificity, while SEU acts as an adaptor between the TFs and LUG (Sridhar *et al.*, 2004, 2006). This allows LUG to

repress target genes using one of two distinct mechanisms: (1) LUG uses its unoccupied intermediate region to bind to the class A HISTONE DEACETYLASE 19 (HDA19), which represses gene expression by removing activating acetyl groups from targeted histones (Fig. 3; Gonzalez *et al.*, 2007). An analogous mechanism is also known for other members of the Gro/TLE super family (Gonzalez *et al.*, 2007; Sridhar *et al.*, 2004). (2) Alternatively, LUG can bind to two members of the mediator complex, CYCLIN-DEPENDENT KINASE 8 (CDK8) and MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 14 (MED14), via both its intermediate region and WD40 domain. This enables LUG to regulate and inhibit the activity of RNA Polymerase II directly, without modifying the chromatin state of the target gene (Gonzalez *et al.*, 2007).

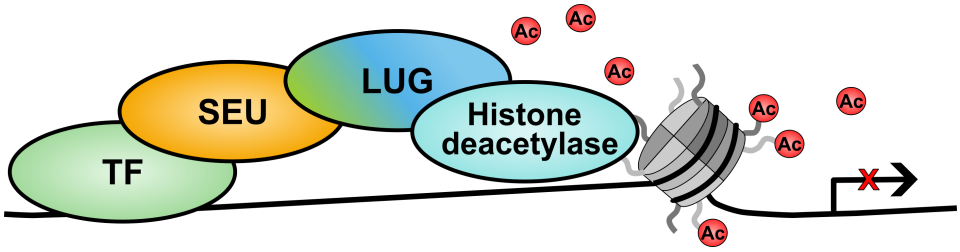


Figure 3: Canonical repression of target genes by the LUG – SEU module. A transcription factor (TF) binds to the adaptor protein SEU, allowing DNA binding and target specificity. SEU binds the LUFS domain (lime) of LUG, which in turn binds to a histone deacetylase that removes acetyl groups from neighbouring histones, thus repressing target gene expression.

1.4.2.3 Further mechanisms of target gene regulation

Besides these two main mechanisms of regulation via the LUG – SEU module, both protein families also exhibit various different gene regulatory mechanisms, usually in combination with different chromatin modulators:

It has previously been suggested that the Gro/TLE family acts as general transcriptional regulators, and not merely as repressors (N. Lee *et al.*, 2015). Indeed, in *A. thaliana*, there are multiple known instances of target gene activation by LUG and LUH: For example, during anther development, LUG has been shown to form an activating regulatory complex with the EAR MOTIF-CONTAINING ADAPTOR PROTEIN (ECAP), which enables target specificity by binding to TFs like BES1/BZR1 HOMOLOG 3 (BEH3), not unlike SEU. Within this complex, LUG then mediates the deposition of activating H3K9ac marks at the target promoter histones H3, via one ore more currently unknown partner(s) (Shi *et al.*, 2024). Another instance is the interaction between LUG and LUH with the mediator complex protein MED25 in plant defense pathways: Similarly to the LUG – SEU module, MED25 acts as an adaptor protein when binding LUG or LUH and recruits the complex to target genes via additional TFs, while LUG/LUH bind to histone remodelers. In the case of LUH, the binding partner is the ARABIDOPSIS HISTONE ACETYLTRANSFERASE OF THE CBP FAMILY 1 (HAC1) which deposits activating H3K9ac marks, while the binding partner for LUG is currently unknown, but causes target gene activation as well (You *et al.*, 2019).

A similar flexibility can also be observed for the SEU family: In *A. thaliana* root tips, SEU is recruited to the *WUSCHEL-RELATED HOMEOBOX 5* (*WOX5*) promoter by SCARECROW (SCR) and directly interacts with the methyltransferase SET DOMAIN GROUP 4 (SDG4) which causes activating H3K4me3 trimethylation and thus promotes *WOX5* expression. This interaction is mediated by the IDR2 of SEU, while the LDB domain is used to bind SCR (Zhai *et al.*, 2020). Likewise, SEU has been shown to interact with the methyltransferase SDG25 as well, which causes H3K4me1 single methylation instead, whose specific function is still debated (Oya *et al.*, 2022; Zhai *et al.*, 2020).

In summary, both the LUG and the SEU families exhibit potential for activating and repressing target gene regulation using multiple different mechanisms of chromatin modification or expression interference. The specific regulatory process used is strongly shaped by the specific interaction partners binding to LUG and SEU within the different gene regulatory pathways.

1.4.3 Protein interactions and regulatory targets of the LUG and SEU families

The LUG – SEU module has been implicated in various developmental and regulatory processes throughout *A. thaliana*, especially during flower development (Fig. 4): LUG and SEU are critical for maintaining proper floral whorl identity by repressing the floral organ identity MADS-box gene *AGAMOUS* (*AG*) in the outer whorls in combination with AP1 and SEP3 (Conner & Liu, 2000; Franks *et al.*, 2002; Sridhar *et al.*, 2006), and by interacting with AP2 to repress the expression of the AP2 antagonistic microRNA miRNA172, thus preventing ectopic expression of either AG or miRNA172 in the outer floral whorls (Grigorova *et al.*, 2011). During gynoecium development, the LUG – SEU module, and possibly LUH, interacts with AINTEGUMENTA (ANT) to ensure ovule initiation, and with INNER NO OUTER (INO) to control the development of the outer ovule integument (Azhakanandam *et al.*, 2008; Bao *et al.*, 2010; Z. Liu *et al.*, 2000; M. K. Simon *et al.*, 2017). Additionally, LUG and SEU play a role in adaxial / abaxial polarity regulation, vascular development, and embryonic shoot apical meristem development and maintenance, most likely by interacting with the abaxial identity promoting YABBYs (Franks *et al.*, 2006; Stahle *et al.*, 2009). Both LUG and SEU are also linked to the auxin signaling pathway, for example via their interaction with the auxin synthesis activating STY1 TF, or through the interaction of SEU with the auxin response factor ETTIN (ETT) (J. E. Lee *et al.*, 2014; Pfluger & Zambryski, 2004; Ståldal *et al.*, 2008).

In addition to their functionality as the LUG – SEU module, both families have been shown to act independently of the module as well: For example, SEU, but not LUG, interacts with a SEP3 – SEEDSTICK (STK) dimer to differentiate the seed abscission zone by regulating lignification during gynoecium development (Azhakanandam *et al.*, 2008; Balanzà *et al.*, 2016; Di Marzo *et al.*, 2022). Furthermore, SEU activity has also been linked to gibberellin signaling during root development (Gong *et al.*, 2016), to the maintenance of root stem

cells (Zhai *et al.*, 2020), to light- and temperature-dependent morphogenesis (Huai *et al.*, 2018), and to hyperosmotic stress response (B. Wang *et al.*, 2022).

Within the LUG family, LUG and LUH share a set of overlapping functions, but also possess distinct regulatory roles: For example, while *lug* mutations show severe flowering defects, no visible effects are present for *luh* flowers. However, homozygous *lug luh* double mutants are embryo lethal, indicating that LUG and LUH share a functional intersection that, to some extent, can be compensated by either homolog (Sitaraman *et al.*, 2008). On the other hand, the expression patterns of LUG and LUH during flower development differ from each other, with LUG being more strongly expressed in the inner whorls, whilst LUH shows a broad, high expression level throughout development (Vega-León *et al.*, 2025). Whereas LUG and the adaptor protein ECAP play an important role during anther development (Shi *et al.*, 2024), LUH has been demonstrated to positively regulate seed growth (Di Marzo *et al.*, 2022), as well as polysaccharide maturation and extrusion (Saez-Aguayo *et al.*, 2013; Walker *et al.*, 2011). Additionally, both LUG and LUH function in various stress response pathways, but often in opposite manners, linking developmental regulation to stress response in plants (Shrestha *et al.*, 2014; Sitaraman *et al.*, 2008). In summary, especially in floral development, the predominant regulation is governed by LUG, while LUH adds an underlying layer of modulation, thus expanding the regulatory possibilities of the LUG family (Vega-León *et al.*, 2025).

In a similar vein, while the specific functions of the three SLKs in *A. thaliana* are only roughly understood at the moment, they have been suggested to share some regulatory overlap with SEU, but are not perfect substitutes for each other (Bao *et al.*, 2010). Especially within stress pathways, functional differences between SEU and the SLKs could be observed: SLK1 and SLK2 can interact with LUH to form a regulatory complex in salt and osmotic stress response, while SLK2 is additionally linked to aluminum resistance in roots, and to glucosinolate synthesis in plant – herbivore interactions via the master regulator MYB51 (Frerigmann *et al.*, 2014; Geng *et al.*, 2017; Shrestha *et al.*, 2014). In contrast, while SEU functions in hyperosmotic stress response, too, its regulation is linked to SEU's ability to form liquid-like nuclear condensates, which the SLKs are unable to do (B. Wang *et al.*, 2022).

To summarize, both the LUG and the SEU families regulate a plethora of different aspects throughout *A. thaliana*, and especially in floral development, both as the LUG – SEU module, and as individual proteins. Their homologs often fulfill redundant functions, but also exhibit divergent regulatory roles due to differing expression patterns and interaction partners. Thus, both families are core hubs not only within the organ developmental network of *A. thaliana*, but also connect phytohormone signaling and stress response to plant sexual development (Fig. 4; J. E. Lee *et al.*, 2014; Shrestha *et al.*, 2014; Vega-León *et al.*, 2025).

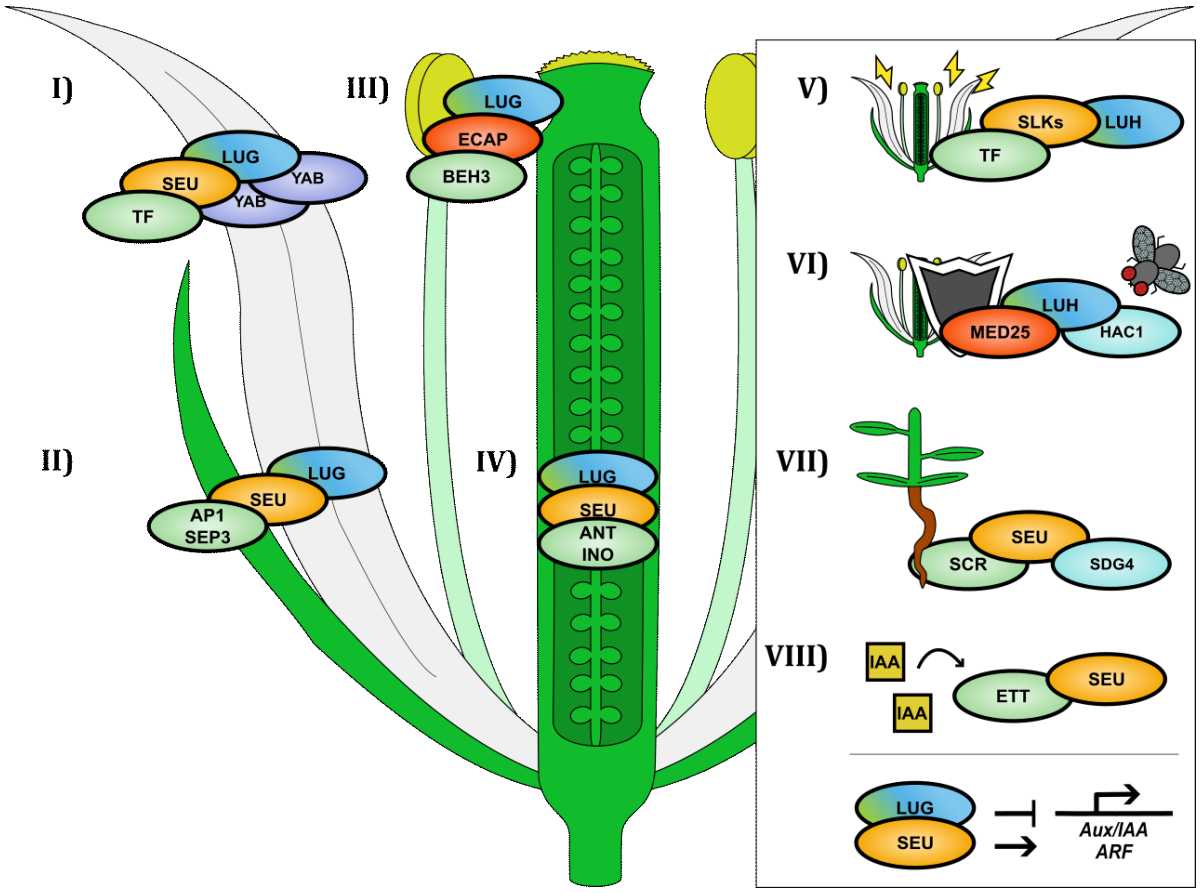


Figure 4: Functions of the LUG and SEU families throughout *A. thaliana* development. Selected protein interactions between LUG (lime – blue) and SEU (orange) family proteins with TFs (light green), histone modifiers (light blue), and other adaptor (red) or further proteins (purple). **I)** Adaxial / abaxial polarity regulation in petals (Franks *et al.*, 2006; Stahle *et al.*, 2009). **II)** Repression of AG in outer floral whorls (Conner & Liu, 2000; Franks *et al.*, 2002; Sridhar *et al.*, 2006). **III)** Regulation of anther development via LUG and ECAP (Shi *et al.*, 2024). **IV)** Ovule development regulation (Azhakanandam *et al.*, 2008; Bao *et al.*, 2010; Z. Liu *et al.*, 2000; M. K. Simon *et al.*, 2017). **V)** LUH and SLKs function in various stress responses (Geng *et al.*, 2017; Shrestha *et al.*, 2014). **VI)** Plant defense regulation by LUG and LUH (You *et al.*, 2019). **VII)** SEU-mediated histone methylation via SDG4 in root tips (Zhai *et al.*, 2020). **VIII)** Interaction with auxin (yellow squares) signaling via ETT, and general regulation of the pathway (J. E. Lee *et al.*, 2014; Pfluger & Zambryski, 2004; Ståldal *et al.*, 2008).

1.4.4 The functions of LUG and SEU homologs beyond *A. thaliana*

Homologs of the LUG and SEU families have been shown to interact in a similar manner in other angiosperms, like in the snapdragon species *Antirrhinum majus*, where they generally regulate organ development, and also specifically interact with homologs of the YABBY family (Navarro *et al.*, 2004), and in the legume *Medicago truncatula*, where LUG homologs positively regulate leaf blade outgrowth and floral development (F. Zhang *et al.*, 2019). In rice, the formation of an interaction complex between LUG, SEU, AP1, and SEP3 homologs has been demonstrated, which regulates many aspects of flower development (Yang *et al.*, 2019).

Phylogenetic analyses revealed that members of the LUG and SEU families were already present in the MRCA of land plants (Pfannebecker *et al.*, 2017), and as such, expression of LUG and SEU homologs can be found in various plant species, like in the hornwort

Anthoceros agrestis, or in *C. richardii* (F.-W. Li *et al.*, 2020; D. B. Marchant *et al.*, 2022). Additionally, condensation of SEU homologs into liquid-like condensates has been shown to occur across land plants, but it remains elusive whether these condensates function in a similar manner in stress responses as they do in *A. thaliana* (B. Wang *et al.*, 2022). However, the origin of the LUG and SEU families beyond the land plant lineage and the exact beginning of their joint developmental regulation is currently unknown.

1.5 The MADS-box transcription factor family

MADS-box genes are a strongly conserved family of developmental control genes that are present in nearly all eukaryotic lineages. They are thought to have originated at least one billion years ago, and encode for TFs containing a „MCM1/AG/DEFICIENS/SRF“ (MADS) domain for target DNA motif binding (Kaufmann *et al.*, 2005; Thangavel & Nayar, 2018). In the green lineage, all MADS-box genes share a common origin, stemming from a *MYCOTE ENHANCER FACTOR-2* (*MEF2*) MADS-box ancestor, and are commonly divided into two families, the M-type and the MIKC group, depending on their domain structure (Qiu *et al.*, 2023; Qiu & Köhler, 2026). The TFs of former group consist of the MADS domain and a variable C-terminus, while the MIKC genes encode for four distinct domains: The MADS (M) domain for DNA binding, the intervening (I) domain for dimerization, a keratin-line (K) domain allowing establishment of protein interactions, and a variable C-terminal (C) domain (Gramzow & Theissen, 2010; Kaufmann *et al.*, 2005; Rümpler *et al.*, 2023). In land plants, the MIKC type can be divided into two subgroups, the MIKC^{C} and the MIKC* MADS-box TFs, which primarily differ in the length of their I-domain (Kaufmann *et al.*, 2005; Kwantes *et al.*, 2012). Up until recently, it had been the predominant thought that the M-type and the MIKC types form two different families of plant MADS-box TFs, with the MIKC^{C} and MIKC* groups being closer related to each other than to M-type. However, novel analyses show that both types originated from a common MIKC precursor in the lineage leading to land plants, with an early divergence of the MIKC^{C} followed by a second divergence of the MIKC* and M-type families (Gramzow & Theissen, 2010; Qiu *et al.*, 2023; Qiu & Köhler, 2026).

During land plant evolution, a strong expansion of the MADS-box family occurred, leading to a high amount of functional diversification within the family (Kaufmann *et al.*, 2005; Qiu *et al.*, 2023). Many extant MADS-box TFs are connected to reproductive regulation, with especially the MIKC^{C} type playing a prominent role in sporophytic reproduction, for example as floral homeotic genes such as *AG*, *SEP3*, or *AP1* in *A. thaliana*, or as regulators of flowering time, meristem identity, and seed development in angiosperms (Folter *et al.*, 2005; Gramzow & Theissen, 2010). In contrast, the MIKC* and M-type genes are suggested to preferably govern the male and female gametophyte, respectively (Qiu & Köhler, 2026). As such, the MIKC* genes of *A. thaliana* are restricted to regulate pollen development and germination, whilst the M-type genes tend to function during the development of the embryo and seed (Gramzow & Theissen, 2010; Qiu & Köhler, 2022; Rümpler *et al.*,

2023; L. Zhang *et al.*, 2024; M.-X. Zhang *et al.*, 2020).

1.6 The mechanism of the auxin signaling pathway in plants

The roles of and genetic regulation by the phytohormone auxin encompass one of the largest signaling and regulatory pathway in land plants, connecting cellular growth, environmental and stress response, plant developmental regulation, and more, using one class of regulatory molecules. There are several kinds of natural and synthetic auxins, with indole-3-acetic acid (IAA) being the most prominent natural auxin, while naphthalene-1-acetic acid (NAA) and 2,4-dichlorophenoxyacetic acid (2,4-D) are often used synthetic auxins (S. Simon & Petrášek, 2011). The canonical auxin signaling pathway using IAA as a signal molecule originated in the MRCA of land plants and consists of three principal components: the TRANSPORT INHIBITOR RESPONSE1 (TIR1)/AUXIN-SIGNALING F-BOX (AFB) auxin sensors, the Aux/IAA transcriptional repressors, and the AUXIN RESPONSE FACTOR (ARF) TF family (Hernández-García *et al.*, 2024; Taylor & Bargmann, 2025). Under low auxin concentrations, ARF-mediated gene expression is blocked by the activity of the Aux/IAA proteins, which bind to the ARFs and recruit the TPL and TPR transcriptional repressors, resulting in chromatin condensation and thus repression of target gene expression. Under high auxin concentrations however, auxin molecules bind to the TIR1/AFB auxin sensors, which mediate polyubiquitylation and degradation of the Aux/IAAs, thus relieving the repression of the ARFs, which in turn activates target gene expression (Taylor & Bargmann, 2025). Additionally, there are several non-canonical regulatory mechanisms known as well, such as the pathway involving SEU and ETT in *A. thaliana* female reproductive regulation (Pfluger & Zambryski, 2004; Simonini *et al.*, 2016; Taylor & Bargmann, 2025). While all three auxins mentioned above are able to bind to TIR1/AFB auxin sensors, there are some mechanistic differences between the different auxin, with especially NAA being able to diffuse through cellular membranes (Delbarre *et al.*, 1996; S. Simon & Petrášek, 2011). Furthermore, auxin can act independently of transcriptional regulation, using a fast response pathway to rapidly adapt to environmental signals: This regulation is governed by the cyclic adenosine monophosphate (cAMP) signal molecule produced by TIF1/AFBs, which causes acidification of the apoplast, leading to rapid cell expansion by allowing cell wall loosening proteins to function optimally. Notably though, both pathways still act in concert, with the transcriptional regulation further modulating this acidic growth, albeit in a slower fashion (Kumar *et al.*, 2025).

Auxin accumulation itself is controlled by multiple processes such as auxin biosynthesis, degradation, and polar transport through the plant. This latter process is mediated by three protein families, namely the AUXIN RESISTANT 1 (AUX)/LIKE-AUX1 (LAX) importer, and the PIN and ATP-BINDING CASSETTE B (ABCB) efflux carrier families. Their tempo-spatial distribution creates a polar auxin transport gradient that is a core part of the auxin regulatory network and allows for the creation of local auxin maxima and minima and thus divergent developmental control. Among the auxin transporters, the PIN

proteins are especially of note, and can be divided into the canonical long PIN proteins localized at the plasma membrane, and the non-canonical short PIN proteins at the endoplasmatic reticulum (Musazade *et al.*, 2025).

Interestingly, recent studies showed that there are three classes of ARF proteins with widely different functions and regulatory mechanisms: While class A ARFs function via the canonical auxin signaling pathway and act as transcriptional activators, class B and class C ARFs are transcriptional repressors instead. Although class B ARFs are known to interact with TPL independently of Aux/IAA, their exact mode of action, as well as many aspects of the class C ARFs in general, remain not fully understood (Lavy *et al.*, 2016; Mutte *et al.*, 2018; Prigge *et al.*, 2025). It has been suggested that class B ARFs might act as an additional layer of genetic regulation by competing with class A ARFs for binding to promoter regions: The strength of their repression is lower than that of Aux/IAA proteins, and knockout of class B ARFs results in overall stronger genetic repression due to increased presence of class A ARFs and subsequent enhanced recruitment of Aux/IAA repressors (Lavy *et al.*, 2016).

In summary, the auxin signaling pathway represents a deeply connected and complexly regulated signaling pathway in the land plant lineage, connecting various aspects of cellular growth, stress response and developmental regulation across whole plants.

1.7 Hypothesis and aim of the dissertation

The various functions of the transcriptional regulators LUG and SEU during development of the *A. thaliana* flower are crucial to ensure the formation of the filial generation, and together with the roles of LUG, SEU, and their homologs in general plant development, plant defense, and stress response, underscore both the interconnection between these different processes, and the central role transcriptional regulators play in proper plant development (Conner & Liu, 2000; Franks *et al.*, 2002; Geng *et al.*, 2017; Shrestha *et al.*, 2014; Sridhar *et al.*, 2004). However, while the functions of LUG and SEU are well studied in *A. thaliana*, many aspects of the evolution and regulation of these families in non-angiosperms are still poorly understood.

This dissertation aims to elucidate several aspects of LUG and SEU homologs within multiple major lineages of land plants, and for streptophyte algae: 1) An extensive phylogenetic reconstruction was performed, tracing the gain and loss of members of the LUG and SEU families not only across land plants, but also in the streptophyte algae, which has not been done previously. This allows for a better understanding about their exact emergence in the green plant lineage, and about their evolution across land plants. Additionally, 2) the evolution and conservation of individual LUG and SEU protein sequences and of their interaction with each other was analyzed, which together with 3) *in vivo* analysis of protein – protein interactions between LUG and SEU homologs, resulted in an expansive study of

the coevolution of these two transcriptional regulators over roughly 800 million years of plant evolution (Bierenbroodspot *et al.*, 2024).

Canonical target gene regulation by LUG and SEU requires both the formation of the LUG – SEU module, and binding to partner proteins for both target specification and to enable target gene regulation (Gonzalez *et al.*, 2007; Sridhar *et al.*, 2004). One of the most important group of genetic targets and interaction partners for the LUG – SEU module in angiosperm flower regulation are the MADS-box genes and TFs, respectively, and especially MIKC^{C} types. Given the presence and conservation of LUG, SEU, and MADS-box TFs across land plant evolution (Kaufmann *et al.*, 2005; F.-W. Li *et al.*, 2020; D. B. Marchant *et al.*, 2022), it was hypothesized that the LUG – SEU – MADS-box interaction during flower development was not an angiosperm-specific novelty, but rather an adaptation of already existing protein networks to support innovative reproductive strategies. Therefore, 4) protein interactions between various MADS-box TFs with LUG or SEU homologs across major green plant lineages were analyzed *in vivo*, giving not only a deeper understanding of the possibility and specificity of these protein interactions, but also helping to trace the evolution of the module into becoming crucial transcriptional regulators of angiosperm flower development.

Besides their role in regulating floral development, LUG and SEU have been implicated in multiple different developmental and stress pathways (Conner & Liu, 2000; Franks *et al.*, 2002; Geng *et al.*, 2017; Shrestha *et al.*, 2014; Sridhar *et al.*, 2004). To analyze if the LUG or SEU families were already connected to these regulatory processes in the MRCA of land plants, and to which extend, 5) *lug* mutants of the extant moss *P. patens* were created and analyzed, giving not only a better understanding about regulatory targets of LUG in bryophytes specifically, but also about the evolution of the connection between LUG and these regulatory pathways throughout land plant evolution in general.

2 Material

2.1 Chemicals

Agarose gel components

- LE Agarose (Biozym Scientific, Hessisch Oldendorf, Germany)
- 6x MassRuler DNA Loading Dye (Thermo Fisher Scientific, Waltham, USA)
- MassRuler DNA Ladder (Thermo Fisher Scientific)

Antibiotics

- Ampicillin trihydrate (100 $\frac{\text{mg}}{\text{ml}}$ stock; SERVA, Heidelberg, Germany)
- Gentamycin sulfate (50 $\frac{\text{mg}}{\text{ml}}$ stock; Carl Roth, Karlsruhe, Germany)
- Kanamycin sulfate (50 $\frac{\text{mg}}{\text{ml}}$ stock; Carl Roth)
- Spectinomycin sulfate (100 $\frac{\text{mg}}{\text{ml}}$ stock in DMSO; Biochrom AG, Berlin, Germany)

Other chemicals

- DanKlorix (bleach; Colgate-Palmolive, Hamburg, Germany)
- Dimethylsulfoxid (DMSO; Sigma-Aldrich, St. Louis, USA)
- DNA from fish sperm (ssDNA; SERVA)
- Ethanol (Carl Roth)
- Isopropanol (Carl Roth)
- Silwet L-77 (Kurt Obermeier, Bad Berleburg, Germany)

2.2 Enzymes

General enzymes

- Gateway LR Clonase II Enzyme Mix (Thermo Fisher Scientific)
- RNase A (Thermo Fisher Scientific)
- T4 DNA Ligase (New England Biolabs, Ipswich, USA)
- Proteinase K (Thermo Fisher Scientific)
- Quick Ligase (New England Biolabs)

- 10x T4 DNA Ligase Reaction Buffer (New England Biolabs)
- 2x Quick Ligase Reaction Buffer (New England Biolabs)

Polymerases

- DreamTaq DNA Polymerase (Thermo Fisher Scientific)
- Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific)
- Phusion High-Fidelity DNA Polymerase (New England Biolabs)

- 10x DreamTaq Green Buffer (Thermo Fisher Scientific)

- 5x Phire Reaction Buffer (Thermo Fisher Scientific)
- 5x Phusion HF Buffer (New England Biolabs)

Restriction enzymes

- BamHI (New England Biolabs)
- BsaI (New England Biolabs)
- KpnI (New England Biolabs)
- NdeI (New England Biolabs)
- NotI (New England Biolabs)
- SalI (New England Biolabs)
- XhoI (New England Biolabs)
- XmaI (New England Biolabs)

- 10x CutSmart Buffer (New England Biolabs)

2.3 Stock solutions

10 mM dNTPs

- 100 µl 100 mM dATP (100 mM stock; Thermo Fisher Scientific)
- 100 µl 100 mM dCTP (100 mM stock; Thermo Fisher Scientific)
- 100 µl 100 mM dGTP (100 mM stock; Thermo Fisher Scientific)
- 100 µl 100 mM dTTP (100 mM stock; Thermo Fisher Scientific)
- 600 µl H_2O

Edwards-Buffer (modified after Edwards *et al.*, 1991)

- 630.4 mg Tris-HCl pH 7.5 (Carl Roth)
- 292.2 mg NaCl (Carl Roth)
- 186.12 mg Disodium salt EDTA x 2 H_2O (AppliChem, Darmstadt, Germany)
- 0.5 ml 20 % SDS
- Ad 20 ml H_2O

100x Fe-EDTA

- 3.73 g Disodium salt EDTA x 2 H_2O (AppliChem)
- 2.78 g FeSO4 x 7 H_2O (Merck, Darmstadt, Germany)
- Ad 1000 ml H_2O
- Autoclave

10x LiAc pH 7.5

- 102.02 g Lithium acetate dihydrate (Acros Organics)
- 900 ml H_2O

- Adjust pH to 7.5
- Ad 1000 H_2O
- Autoclave

10x Macronutrients (Hickok & Warne, 1998)

- 5 g KH_2PO_4 (Carl Roth)
- 1.25 g NH_4NO_3 (Honeywell Fluka, Charlotte, USA)
- 1.2 g $MgSO_4$ x 7 H_2O (Carl Roth)
- 0.26 g $CaCl_2$ x 2 H_2O (Merck)
- Ad 1000 ml H_2O
- Autoclave

Micro elements (ME) stock (Egener *et al.*, 2002; Reski & Abel, 1985; Schween *et al.*, 2003)

- 845 mg $MnSO_4$ x 1 H_2O (Honeywell Fluka)
- 431 mg $ZnSO_4$ x 7 H_2O (AnalytiChem, Duisburg, Germany)
- 309 mg H_3BO_3 (Honeywell Fluka)
- 41.5 mg KJ (Carl Roth)
- 12.1 mg Na_2MoO_4 x 2 H_2O (Merck)
- 1.46 mg $Co(NO_3)_2$ x 6 H_2O (Acros Organics, Waltham, USA)
- 1.25 mg $CuSO_4$ x 5 H_2O (Merck)
- Ad 1000 ml H_2O
- Store at 4 °C

200x Micronutrients (Hickok & Warne, 1998)

- 372 mg H_3BO_3 (Honeywell Fluka)
- 104 mg $ZnSO_4$ x 7 H_2O (AnalytiChem)
- 74 mg $CuSO_4$ x 5 H_2O (Merck)
- 74 mg $(NH_4)_6Mo_7O_{24}$ x 4 H_2O (Honeywell Riedel-de-Haën, Charlotte, USA)
- 50 mg $MnSO_4$ x 1 H_2O (Honeywell Fluka)
- Ad 1000 ml H_2O
- Autoclave

Other solutions

- 5 mM α-Naphtalene acetic acid (NAA)
    - 9.3 mg α-Naphtalene acetic acid (Duchefa Biochemie, Haarlem, The Netherlands)
    - Ad 10 ml H_2O
    - Sterile filtration

- Acetosyringone in 10 mM $MgCl_2$
  - 500 µl 2 M $MgCl_2$
  - 7.5 mg 3',5'- Dimethoxy-4'-hydroxyacetophenone (Acetosyringone; solve in a small amount of 100 % EtOH beforehand; Sigma-Aldrich)
  - Ad 100 ml H_2O

- 1 M 3-Amino-1,2,4-triazole (3-AT)
  - 840.8 mg 3-Amino-1,2,4-triazole (Sigma-Aldrich)
  - Ad 10 ml H_2O

- 5 mM 6-Benzylaminopurine (6-BAP)
  - 11.3 mg 6-Benzylaminopurine (Duchefa Biochemie)
  - Ad 10 ml H_2O
  - Sterile filtration

- 40 % (w/v) D-(+)-Glucose
  - 40 g D-(+)-Glucose (Sigma-Aldrich)
  - Ad 100 ml H_2O
  - Sterile filtration

- 100 $\frac{\mu g}{ml}$ DAPI solution
  - 10 mg DAPI (Carl Roth)
  - Ad 100 ml H_2O

- 75 % EtOH
  - 7.5 ml Ethanol (Carl Roth)
  - 2.5 ml H_2O

- 80 % Glycerol
  - 8 ml Glycerol (Carl Roth)
  - 2 ml H_2O

- 2 M $MgCl_2$
  - 40.66 g $MgCl_2$ x 6 H_2O (Carl Roth)
  - Ad 100 ml H_2O
  - Sterile filtration

- 50 % PEG 4000
  - 50 mg Polyethylene glycol 4000 (Carl Roth)
  - Ad 100 H_2O

- 20 % SDS
  - 20 g Sodium dodecyl sulfate (SDS, Carl Roth)
  - Ad 100 ml H_2O

1x PBS buffer

- 8 g NaCl (Carl Roth)
- 2.68 g Na_2HPO_4 x 7 H_2O (Carl Roth)
- 245 mg KH_2PO_4 (Carl Roth)
- 201.3 mg KCl (Carl Roth)
- Set pH to 7.4
- Autoclave

PEG/LiAc

- 8 ml 50 % PEG 4000 (Carl Roth)
- 1 ml 10x TE
- 1 ml 10x LiAc
- Sterile filtration

10x TE

- 15.76 g Tris-HCl pH 7.5 (Carl Roth)
- 2.92 g Disodium salt EDTA x 2 H_2O (AppliChem)
- 900 ml H_2O
- Adjust pH to 7.5
- Ad 1000 H_2O
- Autoclave

1x TE/1x LiAc

- 1 ml 10x TE
- 1 ml 10x LiAc
- 8 ml H_2O

2.4 Media

2x YT medium (Sambrook *et al.*, 1989)

- 16 g Tryptone/Peptone ex Casein (Carl Roth)
- 10 g Yeast Extract (Carl Roth)
- 5 g NaCl (Carl Roth)
- For plates: 20 g Agar-Agar Kobe I (Carl Roth)
- Ad 1000 ml H_2O
- Autoclave

C-fern growth medium (Hickok & Warne, 1998)

- 100 ml 10x Macronutrients

- 10 ml 100x Fe-EDTA
- 5 ml 200x Micronutrients
- 750 ml H_2O
- 10 g Agar-Agar Kobe I (Carl Roth)
- Set pH to 6
- Ad 1000 ml H_2O
- Autoclave

KNOP-ME (Egener *et al.*, 2002; Reski & Abel, 1985; Schween *et al.*, 2003)

- 1 g $Ca(NO_3)_2$ x 4 H_2O (Carl Roth)
- 250 mg KCl (Carl Roth)
- 250 mg KH_2PO_4 (Carl Roth)
- 250 mg $MgSO_4$ x 7 H_2O (Carl Roth)
- 12.5 mg $FeSO_4$ x 7 H_2O (Merck)
- 10 ml ME stock
- Ad 900 ml H_2O
- Adjust pH to 5.8
- For plates: 12 g Oxoid purified agar (Thermo Fisher Scientific)
- Ad 1000 ml H_2O
- Autoclave
- Optional: Add phytohormone solution
  - 1 ml 5 mM α-Naphtalene acetic acid (NAA)
  - 1 ml 5 mM 6-Benzylaminopurine (6-BAP)

LB medium (modified after Bertani, 1951)

- 20 g LB-Medium-Powder (AppliChem)
- For plates: 15 g Agar-Agar Kobe I (Carl Roth)
- Ad 1000 ml
- Autoclave
- Add 1 µl antibiotic stocks for working concentrations:
  - 100 $\frac{\mu g}{ml}$ Ampicillin trihydrate (SERVA)
  - 50 $\frac{\mu g}{ml}$ Kanamycin sulfate (Carl Roth)

SD -L / -W / -LW dropout media (Clontech, 2010)

- 26.7 g Minimal SD-Base (Takara Bio, Kusatsu, Japan)
- 640 mg -Leu/-Trp DO Supplement (Takara Bio)
- Add according to desired medium:
  - SD -L media: 100 mg L-Tryptophan (Carl Roth)
  - SD -W media: 100 mg L-Leucine (Sigma-Aldrich)
  - SD -LW media: No supplements

- Ad 950 ml H_2O
- Adjust pH to 5.8
- For plates: 14 g Agar-Agar Kobe I (Carl Roth)
- Ad 1000 ml H_2O
- Autoclave

SD -LWH dropout medium (Clontech, 2010)

- 26.7 g Minimal SD-Base (Takara Bio)
- 640 mg -Ade/-His/-Leu/-Trp DO Supplement (Takara Bio)
- 100 mg Adenine hemisulfate salt (Sigma-Aldrich)
- Ad 950 ml H_2O
- Adjust pH to 5.8
- For plates: 14 g Agar-Agar Kobe I (Carl Roth)
- Ad 1000 ml H_2O
- Autoclave
- Optional: 2 ml 1 M 3-Amino-1,2,4-triazole (Sigma-Aldrich)

SOC medium (modified after Hanahan, 1983)

- 20 g Tryptone/Peptone ex Casein (Carl Roth)
- 5 g Yeast Extract (Carl Roth)
- 584.4 mg NaCl (Carl Roth)
- 186.4 mg KCl (Carl Roth)
- Ad 985 ml H_2O
- Adjust pH to 7
- Autoclave
- 10 ml sterile filtered 40 % (w/v) D-(+)-Glucose
- 5 ml sterile filtered 2 M $MgCl_2$

YPAD medium (modified after Hanahan, 1983)

- 20 g Tryptone/Peptone ex casein (Carl Roth)
- 10 g Yeast Extract (Carl Roth)
- 100 mg Adenine hemisulfate salt (Sigma-Aldrich)
- For plates: 20 g Agar-Agar Kobe I (Carl Roth)
- Ad 950 ml H_2O
- Autoclave
- 50 ml sterile filtered 40 % D-(+)-Glucose

2.5 Kits and Services

Kits

- NucleoSpin Gel and PCR Clean-up (MACHEREY-NAGEL, Düren, Germany)
- NucleoSpin Plasmid Easy Pure (MACHEREY-NAGEL)
- NucleoSpin RNA Plant and Fungi (MACHEREY-NAGEL)
- RevertAid H minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific)
- RNase-free DNase Set (Qiagen, Venlo, Netherlands)
- RNeasy Mini Kit (Qiagen)

Services

- DNA Oligo synthesis (Eurofins Genomics, Ebersberg, Germany)
- GeneArt Custom Gene Synthesis (Thermo Fisher Scientific)
- Sanger sequencing (Microsynth Seqlab, Göttingen, Germany)

2.6 Plasmids

Empty plasmids for protein interaction assays

- pENTR4 (Thermo Fisher)
- pGADT7 (Clontech, 2008)
- pGBKT7 (Clontech, 2008)

Gateway-compatible Y2H and BiFC plasmids (Lu *et al.*, 2010)

- pGADT7-GW
- pGBKT7-GW
- pGTQL1211YN
- pGTQL1221YC

Plasmids for *P. patens* mutagenesis (Mallett *et al.*, 2019)

- pMK-Cas9-gate
- pENTR-PpU6P-sgRNA-L1L4
- pENTR-PpU6P-sgRNA-L1R5
- pENTR-PpU6P-sgRNA-L3L2
- pENTR-PpU6P-sgRNA-L5L4
- pENTR-PpU6P-sgRNA-R4R3

A full list of all plasmids used and created during this dissertation can be found in Supplementary Tables 5 to 10.

2.7 Microbial strains and plant material

Bacterial strains

- *Agrobacterium tumefaciens* GV3101 (*Stock-No.* 90; Source unknown)
- *Escherichia coli* DH5α (Thermo Fisher Scientific)
- *Escherichia coli* JM109 (Thermo Fisher Scientific)
- *Escherichia coli* DH10B (Thermo Fisher Scientific)
- *Escherichia coli* DB3.1 (Thermo Fisher Scientific)

Yeast strains

- *Saccharomyces cerevisiae* AH109 (*Stock-No*. 62; Takara Bio)
- *Saccharomyces cerevisiae* Y187 (*Stock-No.* 434; Takara Bio)

Plant species and ecotypes

- *Ceratopteris richardii* Hn-n (Hickok, 1977)
- *Nicotiana benthamiana*
- *Physcomitrium patens* Reute (Hiss *et al.*, 2017)

2.8 Hardware

General Hardware

- Microscope DM5500 B (Leica Microsystems, Wetzlar, Germany)
- Stereo microscope M165 C (Leica Microsystems)
- Cell strainer 100 µm (Carl Roth)
- Filtropur S filters 0.45 µm / 0.2 µm (Sarstedt, Nümbrecht, Germany)
- Precellys 24 Tissue Homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France)
- Nanodrop 2000c (Thermo Fisher)

Plant growth

- Phytochamber AR-41L3 (Percival Scientific, Perry, USA)
- Plant cabinet E 230 (Rubarth Apparate GmbH, Laatzen, Germany)

Protein interaction assays

- Assist Plus Pipetting Robot (Integra, Biebertal, Germany)
- RoToR HDA Colony Pinning Robot (Singer Instruments, Roadwater, UK)
- Singer Plus Plates (Singer Instruments)

2.9 Software

General software

- Cytoscape 3.10.1 (Shannon *et al.*, 2003)
- Citavi 7 (Lumivero, Denver, USA)
- Fiji 2.17.0 (Schindelin *et al.*, 2012)
- Inkscape v1.4
- LaTex 2e (2025) and LyX 2.4
- Leica Application Suite 4.3.0
- SnapGene Viewer 8.0

Sequence analysis, manipulation and visualization

- BioEdit 7.7 (Hall, 1999)
- IQ-TREE2 (Minh *et al.*, 2020)
- MAFFT (Katoh *et al.*, 2002)
- PyMol 3.0 (open source version, Schrödinger, 2010)
- Python 3.12.3 (distribution: Anaconda)

Statistical analysis

- KaKs_Calculator 3.0 (Z. Zhang, 2022)
- R version 4.4.1 (R Core Team, 2022)
- RStudio (Posit team, 2024)
- R software packages
    - ComplexHeatmap (Gu, 2022)
    - ggplot2 (Wickham, 2016)
    - gitter (Wagih & Parts, 2014)
    - pheatmap (Kolde, 2019)
    - poweRlaw (Gillespie, 2015)

Websites used

- AlphaFold 3 (https://alphafoldserver.com/); Abramson *et al.* (2024)
- BLAST (Altschul *et al.*, 1990) and genome databases:
    - ICIPS Garden Blaster (https://dfg-icips.org/icips-blast/); Priyam *et al.* (2019); Roessner *et al.* (2024)
    - MarpolBase (https://marchantia.info/); Tanizawa *et al.* (2025)
    - Phytozome (https://phytozome-next.jgi.doe.gov/); Goodstein *et al.* (2012)
- Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo); Madeira *et al.* (2024)
- CRISPOR (https://crispor.gi.ucsc.edu/); Concordet & Haeussler (2018)
- EMBOSS WATER (https://www.ebi.ac.uk/jdispatcher/psa/emboss_water); Madeira *et al.* (2024); Smith & Waterman (1981)

- NCBI Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi); J. Wang *et al.* (2023)
- PAL2NAL (https://www.bork.embl.de/pal2nal); Suyama *et al.* (2006)
- T_m Calculator (https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html); Thermo Fisher Scientific
- UniProt (https://www.uniprot.org); The UniProt Consortium *et al.* (2025)
- WebLogo (https://weblogo.berkeley.edu/logo.cgi); Crooks *et al.* (2004); Schneider & Stephens (1990)

- All websites were last accessed and functional on **08.09.2025**.

GitHub repository for supplementary scripts:

https://github.com/Garrecht/800-Million-years-of-co-evolution-in-the-green-plant-lineage-the-case-of-LEUNIG-and-SEUSS.git

All scripts can also be found in the „Supplementary scripts“ section.

3 Methods

3.1 General methods

3.1.1 Target DNA amplification

Polymerase chain reaction (PCR) mastermixes were created and PCRs performed using the respective compositions and thermocycler programs described in Table 1, with annealing temperatures of primers being calculated using T_m Calculator (see section 2.9), and elongation times depending on amplicon length. For PCRs using DreamTaq polymerase, elongation times were calculated using the following formula: 60 s for the first two kilobases (Kb) DNA, then 60 s for each additional Kb of length.

Table 1: Mastermix and PCR program for PCRs...

A) ... using DreamTaq DNA Polymerase

10x DreamTaq Buffer	2.0 µl
10 mM dNTPs	0.4 µl
10 µM Forward Primer	1.0 µl
10 µM Reverse Primer	1.0 µl
DreamTaq DNA Polymerase	0.2 µl
Template-DNA	1.0 µl
H_2O	14.4 µl

Denature	95 °C	3 min	
Denature	95 °C	30 s	⎤
Annealing	XX °C	30 s	35 x
Elongation	72 °C	*see text*	⎦
Polishing	72 °C	10 min	
Hold	4 °C	∞	

B) ... using Phusion DNA Polymerase

5x Phusion HF Buffer	4.0 µl
10 mM dNTPs	0.4 µl
10 µM Forward Primer	1.0 µl
10 µM Reverse Primer	1.0 µl
Phusion High-Fidelity DNA Polymerase	0.2 µl
Template-DNA	1.0 µl
H_2O	12.4 µl

Denature	98 °C	30 s	
Denature	98 °C	10 s	⎤
Annealing	XX °C	30 s	35 x
Elongation	72 °C	30 s/Kb	⎦
Polishing	72 °C	10 min	
Hold	4 °C	∞	

Table 1 (cont.)

C) ...using Phire Hot Start II DNA Polymerase

5x Phire Reaction Buffer	4.0 µl
10 mM dNTPs	0.4 µl
10 µM Forward Primer	1.0 µl
10 µM Reverse Primer	1.0 µl
Phire Hot Start II DNA Polymerase	0.4 µl
Template-DNA	1.0 µl
H_2O	12.2 µl

Denature	98 °C	30 s	
Denature	98 °C	5 s	⎤
Annealing	XX °C	5 s	35 x
Elongation	72 °C	15 s/Kb	⎦
Polishing	72 °C	1 min	
Hold	8 °C	∞	

3.1.2 DNA digestion, ligation, and transformation

For DNA digestion, around 1 µg DNA was mixed with 0.5 µl of each used restriction enzyme, 2 µl 10x CutSmart Buffer, and filled to 20 µl with H_2O. The reaction was usually incubated at 37 °C for 15 min to 1 h, after which the digested DNA was applied onto a 1 % agarose gel and run using standard protocol. DNA fragments of the desired length were excised from the agarose gel and purified using the „NucleoSpin Gel and PCR CleanUp“ kit. The concentration of purified fragments was measured via Nanodrop, and the fragments were sequenced and used for further cloning steps.

DNA ligation was performed using either T4 DNA or Quick Ligase, using a molar ratio of vector to insert of 1:3 to 1:5 to enhance ligation efficiency. Vector and insert were mixed with either 2 µl 10x T4 DNA Ligase Reaction Buffer and 1 µl T4 DNA Ligase, or with 10 µl 2x Quick Ligase Reaction Buffer and 1 µl Quick Ligase. Reactions were then filled to 20 µl with H_2O and incubated at room temperature (RT) for 5 to 10 min. Ligations using T4 DNA Ligase were heat inactivated by incubation at 65 °C for 10 min, while no heat inactivation was performed for ligations using Quick Ligase.

Transformation of plasmids was performed via electroporation into electrocompetent *E. coli* DH5α or JM109. Briefly, between 0.5 and 2 µl plasmid were mixed with 50 µl competent cells in an electrical cuvette, and a voltage of 2500 V was applied for 5 ms. Afterwards, the transformed *E. coli* were mixed with 1 ml SOC medium (Hanahan, 1983) and incubated at 37 °C while being oscillated at 180 rounds per minute (rpm) for 1 h. The cells were then streaked on LB medium plates (Bertani, 1951) containing working concentrations of the appropriate antibiotic: 100 $\frac{\mu g}{ml}$ ampicillin or 50 $\frac{\mu g}{ml}$ kanamycin, depending on the plasmid used. The plates were incubated overnight at 37 °C.

3.1.3 LR Gateway cloning

The LR Gateway cloning reaction was based upon the Gateway cloning system from Thermo Fisher Scientific (Invitrogen, 2003) and used for easy cloning of one insert from an entry vector into multiple destination vectors. Generally, a suitable entry vector is created by cloning the insert into a Gateway-compatible vector containing two attL1 and attL2 sites, using restriction cloning as described in section 3.1.2 to place the insert between the attL sites. Afterwards, Gateway LR Clonase II is used to subclone the attL-flanked insert into a compatible destination vector, using this vectors attR1 and attR2 sites to control insert position and orientation. Depending on the vectors used, other attL and attR site variations are also possible, as is performing multiple Gateway reactions simultaneously into one destination vector.

For the Gateway cloning reaction, 50 to 150 ng entry vector and 150 ng destination vector were mixed and filled to 8 µl with TE buffer or H_2O. Then, 2 µl LR Clonase II enzyme mix was added and the reaction mixed and incubated at 25 °C for 1 h. Afterwards, 1 µl Proteinase K was added and the mixture incubated at 37 °C for 10 min to terminate the LR reaction. Transformation into competent *E. coli* was performed as described in in section 3.1.2, however, only *E. coli* DH5α could be used for selection of destination vectors, as this strain is susceptible against the *ccdB* gene present in Gateway-compatible, empty destination vectors, allowing for selection successful Gateway cloning.

3.1.4 Screening, isolation and conservation of cloned plasmids

To screen transformed *E. coli* colonies for the presence of correctly assembled, transformed vectors, colonyPCRs were performed, using DreamTaq polymerase as described in Table 1A. However, instead of using template DNA directly, individual *E. coli* colonies were picked, transferred onto a master plate of LB agar (Bertani, 1951) containing the appropriate antibiotics, and used as template for the PCR. The master plate was incubated at 37 °C overnight, while the colonyPCR products were applied onto a 1 % agarose gel and run using standard protocol.

E. coli colonies of interest were picked from the master plate and used to inoculate 5 ml liquid LB medium each (Bertani, 1951), containing the appropriate antibiotics. These cultures were subsequently incubated at 37 °C and 180 rpm overnight, after which they were used either for plasmid isolation using the kit „NucleoSpin Plasmid Easy Pure“, or for the creation of glycerol stocks by mixing 1 ml liquid overnight culture with 500 µl 80 % glycerol. These stocks were then mixed, flash-frozen in liquid nitrogen, and stored at -70 °C. They can be regrown by streaking a small fraction of the glycerol stock onto LB medium containing the appropriate antibiotics and incubating the plate at 37 °C overnight.

A comprehensive list of all glycerol stocks used and created during this dissertation can be found in Supplementary Tables 5 to 10.

3.2 *In silico* analyses

The analyses described in the sections 3.2.1 and 3.2.2, as well as the investigation of conserved sequence domains of section 3.2.3 were performed by Clemens Rössner and kindly provided for this dissertation.

3.2.1 Identification of *LUG*, *SEU*, and MADS-box family genes

Members of the *LUG* and *SEU* families in *A. thaliana*, *C. richardii*, *K. nitens*, *M. polymorpha* and *P. patens* were identified by BLAST search (Altschul *et al.*, 1990), using either the Phytozome database (Goodstein *et al.*, 2012), MarpolBase (Tanizawa *et al.*, 2025), or the ICIPS Garden Blaster (Priyam *et al.*, 2019; Roessner *et al.*, 2024). Members of the MADS-box family were identified based on published literature: *A. thaliana*: Folter *et al.* (2005); *C. richardii*: Hasebe *et al.* (1998), Q. Huang *et al.* (2014), Kwantes *et al.* (2012), and D. B. Marchant *et al.* (2022); *K. nitens*: Rümpler *et al.* (2023); *M. polymorpha*: Bowman *et al.* (2017) and Zobell *et al.* (2010); *P. patens*: Rensing *et al.* (2008) and Zobell *et al.* (2010).

A list of all major *LUG*, *SEU* and MADS-box genes used for this dissertation can be found in Supplementary Table 4.

3.2.2 Phylogenetic reconstruction of protein families

To analyze the phylogenetic relationships within the extended LUG and SEU families, as well as for members of the SDG family, additional BLAST searches (Altschul *et al.*, 1990) were performed in the Phytozome database (Goodstein *et al.*, 2012) and the ICIPS Garden Blaster (Priyam *et al.*, 2019; Roessner *et al.*, 2024) for the following species: *Adiantum capillus-veneris*, *Amborella trichopoda*, *Anthoceros agrestis*, *Aquilegia coerulea, Arabidopsis lyrata*, *Azolla filiculoides*, *Carica papaya*, *Chara braunii*, *Chlorokybus atmophyticus*, *Ceraton purpureus*, *Cicer arietinum*, *Cleome violacea*, *Cycas panzhihuaensis*, *Diphasiastrum complanatum*, *Eschscholzia californica*, *Ginko biloba*, *Helianthus annuus*, *Lens culinaris*, *Lepidum sativum*, *Mesostigma viride*, *Nymphaea colorata*, *Oryza sativa*, *Penium margaritaceum*, *Phaseolus vulgaris*, *Salvinia cucullata*, *Selaginella moellendorffii*, *Solanum lycopersicum*, *Solanum tuberosum*, *Sphagnum fallax, Sphagnum magellanicum*, *Thuja plicata*, *Vicia faba*, *Vitis vinifera*, and *Zea mays*. The sequences were aligned with MAFFT (Katoh *et al.*, 2002), and phylogenies were generated using IQ-TREE2 with the maximum-likelihood criterion (Minh *et al.*, 2020) and 1000 bootstrap samples. These comprehensive phylogenies were then used to create simplified phylogenies of selected species to better visualize the main phylogenetic clades within each family.

Note: Due to its extensive nature, not all genes used for these phylogenies could be referenced in the supplements. However, all genes and genomes can be found within the listed databases above [State: **08.09.2025**].

3.2.3 Investigating functional protein domains and creation of a LUFS consensus sequence

Conserved domains of LUG- and SEU-like proteins were analyzed using the Conserved Domain Database in the NCBI Conserved Domain Search (J. Wang *et al.*, 2023), and visualized in R using ggplot2 (Wickham, 2016), alongside glutamine residue positions. AlphaFold 3 (Abramson *et al.*, 2024) was used to create structural models of individual proteins, which were used in PyMol 3.0 (Schrödinger, 2010) to identify secondary protein structures. The positions of α-helices and β-sheets with a pIDDT confidence score of ≥ 70 were extracted using python and visualized using R (Scripts 1 and 2).

For the creation of a consensus sequence of the N-terminus of the LUFS domain, the aligned LUG sequences from section 3.2.2 were trimmed by removing vertical areas were a majority of aligned sequences (> 90 %) showed gaps, as well as everything beyond the first 80 amino acids (aa) of the N-terminus. To increase the quality of the consensus sequence, aberrant LUG sequences were removed too, especially homologs consisting mostly of gaps in this region, or those that strongly deviated from the overwhelming majority of sequences. The resulting consensus sequence of 108 LUFS domains was visualized by WebLogo, using standard settings (Crooks *et al.*, 2004; Schneider & Stephens, 1990).

3.2.4 Calculating gene conservation using K_a/K_s ratios

The ratio of non-synonymous (K_a) to synonymous (K_s) substitutions during the evolution of protein families was calculated in relation to protein homologs in either *K. nitens* or *C. richardii*. As a first step, protein sequences of *A. thaliana*, *C. richardii*, *K. nitens*, *M. polymorpha* and *P. patens* were aligned pairwise to the reference sequence using Clustal Omega (Madeira *et al.*, 2024). These protein alignments were used for a codon-alignment of the coding sequences (CDSs), using PAL2NAL with standard settings (Suyama *et al.*, 2006). Both K_a and K_s values, as well as the K_a/K_s ratio of the aligned CDSs were then calculated using KaKs_Calculator 3.0 (Z. Zhang, 2022). The maximum likelihood Goldman-Yang algorithm (GY, Goldman & Yang, 1994) was chosen for the calculation, alongside a standard genetic code. Bonferroni-correction was used for multiple testing correction, and calculations above the corrected p-value threshold were removed. Mean K_a/K_s ratios were computed in R and visualization of results performed using ggplot2 (Wickham, 2016).

To analyze specific functional domains, sequences were split along domain boundaries prior to K_a/K_s analysis: LUG homologs were divided in the LUFS domain, spanning the first 200 aa, an intermediate region, and the WD40 region starting from the first annotated WD40 domain. SEU proteins were split along the borders of the LDB domain, resulting in the 5' IDR1, the LDB domain and the 3' IDR2.

3.2.5 Visualization of interacting protein residues

AlphaFold 3 (Abramson *et al.*, 2024) was used to model the structure of selected LUG – SEU and of SEU – MADS-box protein dimers from *A. thaliana*, *C. richardii*, *Chara braunii*, *K. nitens*, *M. polymorpha*, and *P. patens*. Interacting aa residues were defined as residues with a pIDDT confidence of > 50 within 4.5 Å of a residue with a pIDDT confidence of > 50 of the partner protein chain (see Parvathy *et al.*, 2024). Python scripts were used to extract the position of all interacting residues in the modeled dimers, which were visualized in R using ggplot2 (Wickham, 2016; Scripts 3 and 4); and to visualize interacting residues in the LUFS domain of LUG in PyMol 3.0 (Schrödinger, 2010). For this, the first 160 aa of each tested LUG protein as well as 20 atoms around each interacting SEU residue were extracted as an image and visualized in Inkscape (Script 5).

For each tested LUG homolog, consistently interacting residues in the first 80 aa of the LUFS domain were identified by searching for interacting residues with a pIDDT confidence > 70 in more than half of all analyzed protein dimers for the specific LUG homolog. These residues were marked on the LUFS consensus sequences (section 3.2.3) with blue bars if consistently present across one or two individual LUG homologs, or with green bars if repeatedly present across three to six individual LUG homologs.

3.2.6 Digital transcriptome and proteome analysis

Gene expression data were obtained from the following published transcriptomes: *A. thaliana*: Mergner *et al.* (2020); *C. richardii:* D. B. Marchant *et al.* (2022); *M. polymorpha*: Bowman *et al.* (2017), Briginshaw *et al.* (2022), Frank & Scanlon (2015), Higo *et al.* (2016), Hisanaga *et al.* (2021), & Julca *et al.* (2021); *P. patens:* Lang *et al.* (2018) & Sreedasyam *et al.* (2023). For *A. thaliana*, protein abundance values were obtained from Mergner *et al.* (2020) as well.

The data was visualized using the R packages pheatmap (Kolde, 2019) and ComplexHeatmap (Gu, 2022).

3.3 Yeast two-hybrid assays for the analysis of protein-protein interactions *in vivo*

3.3.1 Selecting suitable interactions of interest

Suitable LUG and SEU proteins for protein interaction analysis via yeast two-hybrid (Y2H) were selected based on phylogenetic relationships and presence in model organisms (see Fig. 5). For *P. patens*, only proteins found in genome annotation v3.3 (Lang *et al.*, 2018) were used, due to annotation v6.1 not yet being available at the time (Bi *et al.*, 2024). MADS-box protein candidates were identified based on previously published literature (see section 3.2.1) and selected by MADS-box type, with focus on both MIKC^C and MIKC*

group proteins. Two M-type MADS-box proteins of *P. patens* were also chosen due to their strong expression in reproductive sporophytes (see Supplementary Fig. 1D).

Two types of protein interactions were analyzed: Firstly, interactions between LUG and SEU proteins with each other, and secondly, interactions between either LUG or SEU with a partner MADS-box protein. Interactions were strictly grouped by and confined to each species used, no cross-species protein interactions were examined.

A list of selected genes and their sequence IDs can be found in Supplementary Table 4.

3.3.2 Growth of plant material

For the generation of C-fern gametophytes, spores of *C. richardii* ecotype Hn-n were surface-sterilized as described by Hickok & Warne (1998): Spores were mixed in a reaction tube with a 1:4 solution of bleach:water, containing trace amounts of Silwet L-77, and incubated for 10 min with occasional mixing by inversion. Afterwards, they were rinsed several times with and then stored in sterile water for two to three days in darkness, to synchronize spore germination („dark start“). Spores were subsequently sown on C-fern growth media plates and grown in a phytochamber under long-day conditions at 200 – 300 $\frac{\mu mol}{m^2 s}$ white light for 16 h, followed by 8 h of darkness, a temperature of 28 °C, and a relative humidity of 85 % (Hickok & Warne, 1998). C-fern sporophytes were cultivated by sowing spores directly onto wet soil and grown under the same conditions.

P. patens ecotype Reute spores were likewise directly sown on soil and grown in a greenhouse at 22 °C with 16 h of 100 $\frac{\mu mol}{m^2 s}$ white light at 22 °C to generate gametophytes. For the cultivation of *P. patens* protonema, axenic *P. patens* Reute plants on KNOP-ME medium plates (Schween *et al.*, 2003) were obtained from cooperation partners in Kaiserslautern. Several of these gametophores were transferred to liquid KNOP-ME medium and grown under occasional stirring for several weeks at constant 60 - 70 $\frac{\mu mol}{m^2 s}$ white light and 25 °C (Hiss *et al.*, 2017; Meyberg *et al.*, 2020).

3.3.3 Cloning of Y2H-compatible plasmids

Synthesized cDNA was used to amplify the CDSs of selected genes from *C. richardii* or *P. patens*. For *LUG* and *SEU* genes, mRNA was isolated from either *C. richardii* sporophytes or from *P. patens* gametophytes using the kit „NucleoSpin RNA Plant and Fungi“. For MADS-box genes, mRNA was isolated using the kits „RNeasy Mini Kit“ and „RNase-free DNase Set“ from *C. richardii* gametophytes and from *P. patens* whole plants, respectively. Additionally, *P. patens* protonema were harvested from liquid culture via cell strainers and used for mRNA isolation with these kits, too. Isolated mRNA was used for cDNA synthesis using the „RevertAid H minus First Strand cDNA Synthesis Kit“, which in turn was used to amplify the CDSs using Phusion polymerase (Table 1B) and the cloning primers listed in Supplementary Table 11.

The PCR-amplified CDSs were then either directly ligated into the Y2H vectors pGADT7 and pGBKT7, using the compatible restriction sites added by the PCR primers (Clontech, 2008), or cloned into a Gateway-reaction compatible pENTR4 vector (Invitrogen, 2003). For the latter, a LR reaction using Gateway LR Clonase II was then used to subclone the CDSs into Gateway-sites containing pGADT7/pGBKT7 variants pGADT7-GW and pGBKT7-GW, respectively (Lu *et al.*, 2010). All cloned plasmids were sequenced to verify the inserts and the open reading frame, using the sequencing primers listed in Supplementary Tables 14 and 15.

For *A. thaliana*, *C. richardii* and *M. polymorpha*, several plasmids containing selected CDSs were obtained from multiple cooperation partners, which are noted in Supplementary Tables 5, 7, and 9. Missing Y2H constructs were cloned as described above, using the gifted plasmids as PCR templates.

Due to technical difficulties and time constraints, the CDSs of *K. nitens* genes and of *P. patens PpLUG*2, *PpLUG*3, *PpSEU1* and *PpSEU2* could not be amplified from cDNA. Therefore, these sequences were synthesized by the GeneArt Custom Gene Synthesis service, using their annotated CDSs as basis. The synthesized sequences were then cloned into the Y2H vectors as described above.

A list of all plasmids cloned and used during this dissertation can be found in the Supplementary Tables 5 – 11.

3.3.4 Transformation of competent yeast cells

The finished pGADT7 and pGBKT7 vectors were transformed into *S. cerevisiae*, using a modified approach based on both the Yeast Protocols Handbook (Clontech, 2008) and the Yeast Transformation Systems 2 User Manual (Clontech, 2010). Two different types of transformations were performed, either single transformation of one plasmid at a time, or co-transformation of one pGADT7 and one pGBKT7 plasmid simultaneously. For single transformations, *S. cerevisiae* strains AH109 or Y187 were transformed with either pGADT7 or pGBKT7 plasmids, respectively, while co-transformations were performed in AH109.

As a preliminary step, both *S. cerevisiae* strains were plated on solid YPAD media plates from glycerol stocks (see section 2.7) and incubated for three days at 30 °C. These colonies could then be used for yeast transformations for roughly four weeks.

To perform the transformation, two to three yeast colonies were picked from the appropriate plates and used to inoculate 25 ml YPAD medium (Hanahan, 1983) each as starter cultures that were incubated overnight at 30 °C and 180 rpm. The next day, the OD_{600} of these cultures was measured to determine cell density, and one culture with a suitable OD_{600} between two and five was chosen. 15 to 25 ml of the selected overnight culture

were used to inoculate 300 ml YPAD medium, which was then incubated at 30 °C and 180 rpm for 2 to 3 h, until its OD_{600} reached 0.4 to 0.6. At this point, the yeast cells were harvested by centrifugation at 1000 g and RT for 5 min. The supernatant was discarded, while the pellets were resuspended in 1x TE, pooled, and centrifuged again at 1000 g and RT for 5 min. The resulting yeast pellet was resuspended in 0.5 to 1.5 ml 1x TE/1x LiAc, depending on the amount of transformations performed. Meanwhile, ssDNA was denatured by boiling at 99 °C for 5 min and immediately put on ice, to act as carrier DNA during the transformation.

For each transformation, the following components were mixed in a chilled reaction tube:

- 50 µl denatured ssDNA
- 1 µg plasmid (either pGADT7 or pGBKT7 plasmid for single transformation, or both for co-transformation)
- 50 µl yeast cells in 1x TE/1x LiAc
- 600 µl PEG/LiAc

The components were mixed and incubated at 30 °C and 50 rpm for 30 min. Afterwards, 70 µl DMSO were added and the reactions mixed by inverting. A heat shock was performed by incubating the tubes in a 42 °C water bath for 15 min, with occasional mixing. The tubes were subsequently placed on ice for 2 min, after which the cells were pelleted by using the „pulse“ setting of a centrifuge for 5 s. The supernatants were discarded and cells resuspended in 1x TE each and streaked onto the appropriate SD dropout media:

- Single pGADT7 transformation: SD -L dropout medium
- Single pGBKT7 transformation: SD -W dropout medium
- Double pGADT7 and pGBKT7 transformation: SD -LW dropout medium

Media plates were incubated at 30 °C for three days. For easier handling, master plates were created by picking and transferring transformed colonies onto fresh SD dropout medium, which were incubated at 30 °C for three days. Usually, two or three replicates of each master plate were created.

Single transformations were performed for most *C. richardii* and for *M. polymorpha* and *P. patens* interactions, while co-transformations were performed for *A. thaliana*, *K. nitens* and the remaining *C. richardii* interactions, due to problems during the mating process for these species.

3.3.5 Robot-assisted Y2H assay

The Y2H assay was based on the Matchmaker Gold Yeast Two-Hybrid system (Clontech, 2008) and performed as described by Luck *et al.* (2020) and Striebinger *et al.* (2013) on „Singer Plus Plates“. For single transformations, mating of yeast strains had to be performed first, while co-transformed yeast colonies could be used directly for the assay.

Mating of yeast cells was performed by picking transformed colonies from the master

plates and resuspending them to an OD_{600} of 0.5 with H_2O. For each interaction, the appropriate pGADT7 and pGBKT7 plasmids were combined by either plating 4 µl of each cellular suspension on top of each other on YPAD medium plates and incubating them at 30 °C for two to three days (Hanahan, 1983); or by mixing 5 µl of each suspension in 100 µl liquid YPAD medium and incubating them overnight at 30 °C and 30 rpm. Both variations were performed using the „Assist Plus Pipetting Robot".

To select mated yeast cells containing both plasmids, they were transferred onto SD -LW dropout plates, either by replica plating using the „RoToR HDA Colony Pinning Robot", or by transferring 5 µl of the liquid YPAD medium using the „Assist Plus Pipetting Robot". These plates were then incubated at 30 °C for three days, resulting in the growth of yeast colonies containing both pGADT7 and pGBKT7 plasmids.

The Y2H assay was performed in a 96 spot format with three replicates each. For each 96 spot plate, a schematic was created depicting colony position and plasmids tested. Both mated and co-transformed colonies were placed into the 96 spot format on SD -LW plates via the „Assist Plus Pipetting Robot", which were then incubated at 30 °C for three days. Afterwards, the „RoToR HDA Colony Pinning Robot" was used to transfer each 96 spot plate onto three assay medium plates: One SD -LW plate as control, one SD -LWH triple dropout plate, and one SD -LWH plate containing 2 mM of the competitive inhibitor 3-AT, which inhibits expression of the yeast *HIS3* gene at low levels (Clontech, 2008). The assay plates were incubated at 30 °C for three days.

As negative protein controls, all proteins were paired with both empty pGADT7 and with empty pGBKT7 plasmids. To control the assay itself, empty pGADT7 and pGBKT7 plasmids were paired as negative control, while the positive control was constituted of the proteins DEFICIENS1 (DEF1) and SEIRANA (SEI; previously EScaGLO) from *Eschscholzia californica* (Lange *et al.*, 2013).

3.3.6 Documentation and analysis of results

The Y2H assay plates were documented by taking pictures of the plates from a fixed distance using a mounted camera. The images were then cropped using Fiji (Schindelin *et al.*, 2012), and colony sizes were measured using the R package gitter (Wagih & Parts, 2014; Scripts 6 and 7). In case of measurement failure, caused by either background noise or color differences within colonies, Fiji was used to convert images to grayscale, after which a rolling ball background subtraction using a 50 pixel ball radius was performed. Then, a median filter of two pixels was applied, followed by the Fiji Lookup Table „Fire". These transformations allowed a correct, noise free gitter analysis, without negatively affecting analysis precision.

Measured spots with a size below 900 pixels on the SD -LW control plates were discarded as absent colonies or artifacts created by the plating robot. This cutoff value was deter-

mined by analyzing the mean size of actual colonies, which was considerably larger than 900 pixels, while most visual artifacts fell well below this value. The size of the colonies grown on the SD -LWH plates were divided by the size of the identical clones on the SD -LW control plates, to calculate the relative colony growth on the assay plates for each individual colony / interaction tested. Values above one were capped at one. Colonies with a relative growth rate of 0.5 or above were considered to contain interacting proteins. See Supplementary Fig. 2 for a graphical overview of the Y2H analysis.

Proteins where the pGBKT7 plasmid showed protein interaction when paired with empty pGADT7 were deemed autoactivating proteins; and all interactions with autoactivating proteins were removed from the results. Additionally, interactions that were only observed in one of three replicates were also removed. The resulting protein interaction networks were visualized using Cytoscape 3.10.1 (Shannon *et al.*, 2003).

3.4 Verification of selected protein interactions using bimolecular fluorescence complementation

Due to the large scale of the performed Y2H analysis, only selected interactions could be verified via second, independent means: For *A. thaliana*, the observed results were cross-referenced with already published literature, especially with Y2H assays described in independent publications (see Shrestha *et al.*, 2014; Sitaraman *et al.*, 2008; Stahle *et al.*, 2009). *M. polymorpha* results could be verified by Co-Immunoprecipitation (CoIP) and by bimolecular fluorescence complementation (BiFC) assays performed by Q. Li (2022). Finally, for interactions between *C. richardii* LUG and SEU proteins, BiFC assays were performed in *Nicotiana benthamiana* (tobacco) leaves:

3.4.1 Creation of BiFC-compatible plasmids

The *C. richardii* cDNA created in section 3.3.3 was used to amplify the CDSs of *CrLUG*, *CrSEU1*, *CrSEU2*, *CrSEU3* and *CrSEU4*, using the cloning primers listed in Supplementary Table 12. The amplified CDSs were then firstly cloned into the Gateway-reaction compatible pENTR4 vector (Invitrogen, 2003), which was then used to subclone the CDSs into the BiFC plasmids pGTQL1211YN and pGTQL1221YC (Lu *et al.*, 2010) using Gateway LR Clonase II. These plasmids contain the N-terminal or the C-terminal domain of yellow fluorescent protein (YFP), respectively, at the 3' end of the open reading frame, thus requiring the CDSs to lack a stop codon. The cloned plasmids were sequenced to verify the inserts and the open reading frame, using the sequencing primers listed in Supplementary Table 15.

An overview of the cloned BiFC plasmids can be found in Supplementary Table 7.

3.4.2 *N. benthamiana* growth and transformation

Seeds of *N. benthamiana* were sowed on soil and grown in a greenhouse at long-day conditions at 22 °C with 16 h of 100 $\frac{\mu mol}{m^2 s}$ white light for around four weeks. Their leaves were then infiltrated by transformed *A. tumefaciens* GV3101 containing the desired plasmids.

The BiFC assay was performed as described by Lange *et al.* (2013): To prepare the infiltration, electrocompetent *A. tumefaciens* GV3101 cells were transformed with the BiFC plasmids using electroporation (see section 3.1.2). The transformed cells were then streaked on 2x YT medium plates (Sambrook *et al.*, 1989) containing 50 $\frac{\mu g}{ml}$ gentamycin and 50 $\frac{\mu g}{ml}$ kanamycin, and incubated at 28 °C for three days. Additionally, helper plasmid p19 from the tomato bushy stunt virus was also transformed, which suppresses RNA silencing in host cells (Lange *et al.*, 2013; Walter *et al.*, 2004). Afterwards, the resulting colonies were used to inoculate 10 ml liquid 2x YT medium each and incubated at 28 °C and 180 rpm overnight. For every three BiFC assays performed, one 10 ml p19 culture was created. The next day, the OD_{600} of the individual cultures was measured and the cultures harvested by centrifugation at 2000 rpm for 15 min. The supernatants were removed, and the pellets resuspended in 10 mM $MgCl_2$ containing acetosyringone, to an OD_{600} of 1.2; except for the p19 cultures, which were set to an OD_{600} of 2.4. The resuspended cells were incubated at RT for 2 h and then mixed together using the following formula:

- 250 µl p19
- 500 µl pGTQL1211YN plasmid
- 500 µl pGTQL1221YC plasmid

These mixtures were used to infiltrate *N. benthamiana* leaves by filling them into needleless syringes, placing them on the abaxial side of the leaves, and pushing the *A. tumefaciens* solution into the leaves. For each tested combination, several independent leaf infiltrations were performed. The BiFC assay and the tobacco plants were then incubated for two days at RT.

Empty pGTQL1211YN and pGTQL1221YC plasmids were used as negative control, and pMLBART vectors expressing YC/YN-tagged 35S::bZIP63 were used as positive controls (Lange *et al.*, 2013; Walter *et al.*, 2004), which where transformed onto 2x YT medium containing 50 $\frac{\mu g}{ml}$ gentamycin and 100 $\frac{\mu g}{ml}$ spectinomycin (see Supplementary Table 6).

3.4.3 Imaging and analysis of BiFC results

For visualization of the BiFC assays, small leaf discs were excised from the tobacco plants after two days. The discs were washed in 1x PBS buffer and stained with 100 $\frac{\mu g}{ml}$ DAPI solution for 5 min. Afterwards, the DAPI solution was removed and the discs washed several times with 1x PBS buffer. The stained discs were observed and documented using a Leica DCM5500 fluorescence microscope, using filter blocks L5 for YFP fluorescence and

A4 for DAPI stains. Pictures of the observed fluorescence were taken and visualized using Fiji (Schindelin *et al.*, 2012): Firstly, the pictures were cropped and split along the colour channels. Secondly, the green channel containing the YFP fluorescence and the blue channel of the respective DAPI stain were merged, resulting in a composite picture containing both fluorescences (Script 8). These pictures were then used to assess the presence or absence of protein interaction.

3.5 CRISPR/Cas9-mediated mutagenesis of *P. patens*

The CRISPR/Cas9-mediated mutagenesis was performed for the *P. patens LUG* and *SEU* homologs *PpLUG1*, *PpLUG2*, *PpLUG3*, *PpLUG4*, *PpSEU1,* and *PpSEU2* in double (*pplug1 pplug2*; *pplug3 pplug4*; *ppseu1 ppseu2*), quadruple (*pplug1 pplug2 pplug3 pplug4*) and sextuple (*pplug1 pplug2 pplug3 pplug4 ppseu1 ppseu2*) configurations. Due to time constraints and technical difficulties, only the *pplug3 pplug4* double mutant line could be analyzed for this dissertation. To keep the thesis concise, only work relating to this mutant line will be described in detail.

3.5.1 Creation of CRISPR/Cas9 vectors targeting *P. patens LUG* and *SEU* genes

Protospacer for mutagenesis of the selected *PpLUG* and *PpSEU* genes were designed using CRISPOR (Concordet & Haeussler, 2018). „Physcomitrium patens - Phascum patens - NCBI GCF_000002425.4 (Phypa V3)“ was set as reference genome to screen for off-targets, and „20bp-NGG - Sp Cas9, SpCas9-HF1, eSpCas9 1.1“ as protospacer adjacent motif (PAM). The exons plus ∼15 bp of the surrounding intron regions were loaded into the tool, to search for exon-located protospacer motifs. From the resulting lists of possible guide sequences, two for each gene that fulfilled the following criteria were picked:

- No possible off-target sites with three or fewer mismatches present.
- No possible off-targets present where all mismatches are distal to PAM site, beyond 12 bp away (Concordet & Haeussler, 2018).
- Not deemed „inefficient“ by CRISPOR (Concordet & Haeussler, 2018).
- As close to the 5’ end of the gene as possible.
- Sequences that can affect multiple target genes at once due to sequence similarity.
    - Highly similar gene pairs:
        - *PpLUG1 – PpLUG2*: similarity: 80.9 %, gaps: 18.8 %
        - *PpLUG3 – PpLUG4*: similarity: 96.7 %, gaps: 2.1 %
        - *PpSEU1 – PpSEU2*: similarity: 78.6 %, gaps: 21.4 %

Similarity score between highly similar gene pairs was calculated by pairwise alignment of CDSs, using a Smith–Waterman algorithm with a gap penalty of 10.0 and an extension penalty of 0.5 (Madeira *et al.*, 2024; Smith & Waterman, 1981).

For each selected gene, two protospacers were chosen: One that was specific for that gene alone, and one that also targeted the highly similar partner gene. Complementary primers

with added adapter sequences for cloning of the protospacer into the CRISPR-vectors were created (see Supplementary Table 13) and denatured at 98 °C for 3 min. The temperature was then lowered slowly to the melting temperatures of the primer pairs, where they were annealed for 10 min. Finally, the temperature was slowly reduced to RT to obtain usable protospacers.

The used CRISPR/Cas9 system was based on Mallett *et al.* (2019): The three entry vectors pENTR-PpU6P-sgRNA-L1L4, pENTR-PpU6P-sgRNA-R4R3 and pENTR-PpU6P-sgRNA-L3-L2 were digested using BsaI, as described in section 3.1.2. Each of the annealed protospacers were ligated into one of the entry vectors, which were then used in a multiplex Gateway reaction, using Gateway LR Clonase II to subclone all three protospacers from the entry vectors into destination vector pMK-Cas9-gate, resulting in a multiplex Cas9 vector containing *PpLUG3* and *PpLUG4* protospacer, each flanked by a PpU6 promoter and the sgRNA scaffold sequence.

A list of all cloned CRISPR plasmids can be found in Supplementary Tables 6 and 10. Plasmids were sequenced with the primers listed in Supplementary Table 15.

3.5.2 *P. patens* transformation and culture maintenance

Transformation of *P. patens* Reute and selection of transgenic lines was kindly performed by the lab of Prof. Müller-Schüssele in Kaiserslautern, according to a procedure modified after Hohe *et al.* (2004) & Schween *et al.* (2003).

The resulting axenic *P. patens* mutant lines and wildtype (WT) control plants were grown on KNOP-ME medium at 16 h of 60 – 70 $\frac{\mu mol}{m^2 s}$ white light and 8 h of darkness at 22 °C for optimal growth (Hiss *et al.*, 2017; Meyberg *et al.*, 2020). The mutant lines were continuously maintained by transferring material onto fresh KNOP-ME plates under sterile conditions every few months, and older plates were stored in darkness at 4 °C to inhibit further growth, while keeping the lines alive (Egener *et al.*, 2002; Reski & Abel, 1985; Schween *et al.*, 2003).

3.5.3 Genotyping of CRISPR mutant lines

For genotyping of *P. patens* mutant lines, DNA was isolated from *P. patens* plants by picking material (usually two to three gametophores) from the KNOP-ME plates and transferring them into screw-cap tubes containing a few ceramic beads each. The material was disrupted in a tissue homogenizer twice at 4.000 rpm for 90 s each. Edwards-buffer was added to the tubes, and the samples were centrifuged at 17.000 g for 2 min. 150 µl of the supernatants were transferred into fresh tubes, mixed with 150 µl isopropanol, and incubated at RT for 10 min. Afterwards, the samples were centrifuged for 10 min at 17.000 g and the supernatants removed. 200 µl cold 75 % EtOH were added and the samples again

centrifuged at 17.000 g for 5 min. The EtOH was removed and the DNA pellets dried at 60 °C for 5 min. Finally, the pellets were resuspended in 20 µl H_2O containing RNase A and incubated at 60 °C for 2 min to obtain DNA samples.

Genotyping of mutant loci was performed using either DreamTaq DNA polymerase or Phire Hot Start II DNA Polymerase, as described in Table 1. The primers used are described in Supplementary Table 16 and were usually located around 200 bp up- and downstream of the CRISPR target site. Generally, PCRs were either performed for individual target sites, or spanning both CRISPR target sites at once.

Amplified PCR fragments were screened on 1 % agarose gels for detection of large deletions, encompassing the whole region between the two CRISPR target sites. Alternatively, 5 µl PCR fragments of mutant lines were mixed with 5 µl fragments amplified from wildtype using identical primers. The mixtures were boiled at 98 °C for 5 min and slowly cooled to RT, which allowed the formation of homo- and heteroduplexes between mutant and wildtype DNA fragments. These duplexes were then applied onto 4 % agarose gels and run for around 2 h, which permitted visualization of ≥ 2 bp mismatches between mutant and wildtype DNA fragments (Bhattacharya & van Meir, 2019).

DNA fragments were purified using the „NucleoSpin Gel and PCR Clean-up“ kit and sequenced using Sanger sequencing. For lines with verified mutations, all other targeted CRISPR sites were sequenced as well, and a visual alignment of the mutant sites was created.

Additionally, AlphaFold 3 models of the PpLUG3 protein and of the resulting protein of the L3_2 deletion line were created and used to create a comparison of the first 130 aa of both proteins in PyMol 3.0 (Abramson *et al.*, 2024; Schrödinger, 2010).

3.5.4 Phenoytping of the *P. patens* reproductive cycle and phytohormone signaling defects

Due to time constraints, only the reproductive cycle of *P. patens* and the influence of the phytohormones auxin and cytokinin on eleven individual *pplug3 pplug4* mutants and on WT plants were phenotyped in greater detail:

For the analysis of the reproductive cycle, singular gametophores were transferred onto fresh KNOP-ME medium and grown for four weeks under long-day conditions at 16 h 60 – 70 $\frac{\mu mol}{m^2 s}$ white light and 22 °C (Hiss *et al.*, 2017; Meyberg *et al.*, 2020). Both the Leica DM5500 microscope as well as the M165 C stereo microscope were used to measure the sizes of the individual *P. patens* gametophytes, which were calculated using Fiji (Schindelin *et al.*, 2012). To determine significant differences, Student's *t*-test was performed in R, using the Benjamini–Hochberg procedure to correct for multiple testing (Benjamini & Hochberg, 1995; Student, 1908). Besides colony sizes, the following characteristics were recorded and categorized after both four weeks and eight weeks of growth

under long-day conditions: Protonema types observable, protonema growth and distribution, as well as protonema cell structure and chloroplast distribution. For protonema growth, the length and width of 20 to 40 individual chloronema cells of each mutant line were measured in Fiji and visualized in R using ggplot2 (Wickham, 2016), excluding terminal cells. Pairwise Student's *t*-test with Benjamini–Hochberg correction was used to determine significant differences in cell length and width (Benjamini & Hochberg, 1995; Student, 1908). Gametophore growth and development was categorized by presence, gametophore coverage and development, leaflet shape and cell structure and chloroplast distribution.

The formation of gametangia was induced by transferring four week old *P. patens* plants to short-day conditions at 15 °C, with 8 h of 20 – 30 $\frac{\mu mol}{m^2 s}$ white light and 16 h of darkness. Gametangia formation was documented after three weeks, and the plants flooded with sterilized tap water to induce fertilization. Growth under short-day conditions was continued for four weeks, after which the rate of sporophyte formation was measured (Hiss *et al.*, 2017; Meyberg *et al.*, 2020): The total amount of gametophores carrying no sporophytes, green sporophytes or brown sporophytes were counted and a pairwise Fisher's exact test with Benjamini–Hochberg correction performed to analyze statistical differences between WT and mutant line sporophyte distributions (Benjamini & Hochberg, 1995; Fisher, 1970). Additionally, the presence and structure of spores was documented.

The influence of the phytohormones auxin and cytokinin on the growth and development of the *P. patens* mutants was determined by growing them on KNOP-ME medium containing either 5 µM auxin (NAA) or 5 µM Cytokinin (6-BAP) under long-day conditions for four weeks (Ashton *et al.*, 1979). Afterwards, the effects of the phytohormones were documented using a Leica M165 C stereo microscope and a DM5500 microscope, as described above. The following characteristics were documented: Overall plant habit, gametophyte growth, and protonema characteristics such as protonema types, growth and distribution. For mutants grown on exogenous auxin, chloronema cell length and width was measured as described above, and compared to chloronema sizes of WT lines or to those grown on normal growth medium, using Student's *t*-test and Benjamini-Hochberg multiple testing correction (Benjamini & Hochberg, 1995; Student, 1908).

4 Results

The results of the sections 4.1 to 4.4 are largely parts of the submitted, but, as of the **08.04.2026**, still unpublished paper „800 million years of co-evolution in the green plant lineage – the case of LEUNIG and SEUSS transcriptional co-regulators“ (Garrecht *et al.*, 2026). Contributions of the individual coauthors are noted or cited at the appropriate sections. The *P. patens pplug* mutant lines described in section 4.5 have neither been published nor submitted so far.

4.1 Homologs of LUG and SEU can be found in all major land plant lineages and in streptophyte algae

The analyses described in the following section have been performed by Clemens Rössner for Garrecht *et al.* (2026), and were kindly provided for this dissertation as well.

While the phylogenetic relationships of LUG and SEU homologs are well known in angiosperms (Pfannebecker *et al.*, 2017), considerably less is known both about their relation in lycophytes and bryophytes, and whether their families emerged in the lineage leading to land plants, or if they predate it. Therefore, representative species of all major land plant and streptophyte algae lineages were picked for an expansive phylogenetic reconstruction of LUG (Supplementary Fig. 3) and SEU homologs (Supplementary Fig. 4) throughout the evolution of the streptophytes. These comprehensive phylogenies were then simplified and used for an analysis of conserved protein domains (Fig. 5). Both families possess at least one homolog in all lineages examined, as well as similar patterns of conserved protein domains and secondary structures.

Land plant LUG homologs represent a sister group to algae LUG homologs, and consist of the LUG-clade, the LUH-clade, and an unresolved polytomy (Fig. 5A). Both the LUG- and LUH-clades are monophyletic clades with 100 % bootstrap support each, and only contain homologs from seed plants, which is consistent with previously described angiosperm LUG- and LUH-clades (Pfannebecker *et al.*, 2017). However, the relationship of bryophyte, lycophyte, and monilophyte LUG homologs in respect to the LUG- and LUH-clades couldn’t be resolved fully.

All analyzed LUG homologs contain between two to five annotated, conserved WD40 domains at the C-terminal end of the proteins. While the propeller-like structure created by WD40 domains usually consists of seven individual WD40 domains, each with a set of four β-sheets, they have been reported to be highly divergent and thus difficult to annotate using sequence conservation (Xu & Min, 2011). Indeed, when looking at the identified β-sheets alone, most LUG homologs contain seven sets of four β-sheets with a high pIDDT confidence above 70 at their C-terminus, revealing a structural, but not a sequence conservation of the WD40 propeller. Furthermore, all LUG homologs are rich in glutamine, especially near the N-terminus, where large stretches consisting entirely out of Q residues

can be found. Finally, the LisH domain, which consists of two short α-helices, can be found at the N-terminus of most land plant LUG homologs, usually followed by a long α-helix. Notably, the LUG homolog of *K. nitens* is the only algae LUG protein containing this domain. A separate BLAST using only the *K. nitens* LUFS domain as query wasn't able to identify any additional algae LUFS domains as well. Generally, missing LisH domains in land plant and algae proteins could be explained by improper genome annotation, especially considering the conservation of the LisH motif across eukaryotes (Gerlitz *et al.*, 2005; Z. Liu & Karmarkar, 2008), and the lower quality genomes present for most algae (see viridiplantae BUSCO scores of used genomes, Priyam *et al.*, 2019; Roessner *et al.*, 2024). The conservation of the whole LUFS domain could not be checked, as it is currently not present in the Conserved Domain Database used for this analysis (J. Wang *et al.*, 2023).

The phylogeny of the SEU family can be resolved into three larger monophyletic clades and an unresolved polytomy (Fig. 5B): Both the SEU-clade and the SLK-clade contain only seed plant homologs with 84 % and 100 % bootstrap support, respectively, while the homologs of monilophytes and bryophytes belong to the polytomy. Both clades and the polytomy form a phylum underpinned by 79 % bootstrap support, which is in line with preceding phylogenies (Pfannebecker *et al.*, 2017), and several algae SEU homologs are 96 % supported sisters to them. Finally, there is the SEU/SLK-sister-clade with 97 % bootstrap support, which represents a sister lineage to all other SEU homologs and incorporates sequences of all major land plant lineages besides angiosperms, suggesting a loss of their members in flowering plants.

A centrally located LDB domain is present for all analyzed SEU homologs, usually folding into several β-sheets flanked by α-helices, with the notable exception of PpSEU2 and a *S. moellendorffii* homolog. For some homologs, additional high-confidence α-helices beyond the LDB domain region can be found, too, but large parts of the N- and C- termini lack fixed secondary protein structures in confirmation with their roles as IDRs (B. Wang *et al.*, 2022). Similar to the LUG homologs, all SEU proteins are glutamine rich, with the majority of Q residues usually being located N-terminal of the LDB domain, especially in SEU- and SLK-clade members.

In summary, these phylogenetic reconstructions and domain analyses demonstrate both the well supported presence of LUG and SEU homologs in not only all major land plant species, but also in streptophyte algae, as well as the conservation of functional domains and secondary structures across most members of these families, with the notable exception of the LisH domain. However, for both families, algae sequences are confined to sister clades of land plant sequences.

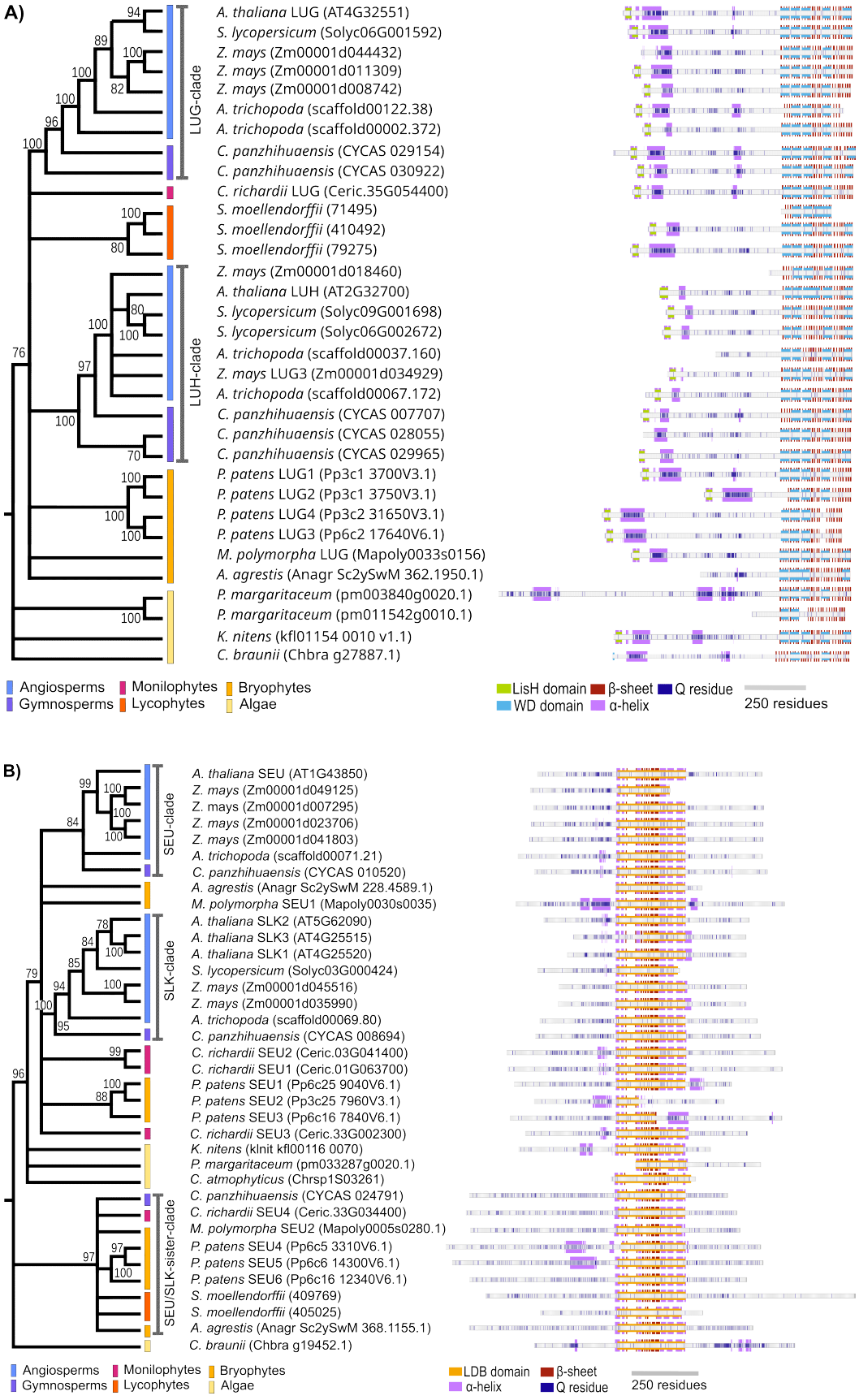


Figure 5: Simplified phylogenies and conserved protein domains of LUG and SEU family proteins. Reduced phylogenies of **A)** LUG, LUH, and LUG-sister proteins, and **B)** SEU, SLK, and SEU/SLK-sister proteins, with bootstrap-support shown on selected nodes. Branches with bootstrap support below 75 were collapsed. The position of conserved domains (LisH in lime, WD40 in light blue, and LDB in orange), high-confidence secondary protein structures (α-helices in purple, β-sheets in red), and Q residues (dark blue) within the proteins are shown on the right.

4.2 *LUG* and *SEU* functional domain regions are under notable purifying selection across plant evolution

To further examine the observed conservation of LUG and SEU domains and of secondary protein structures, the selective pressure extorted upon their genes during evolution was calculated by determining the ratio between non-synonymous (K_a) to synonymous (K_s) substitutions between distant homologs. A K_a/K_s ratio < 1 indicates a purifying selection removing deleterious alleles, while ratios > 1 depict adaptive selection of genes. Sequences with values ≈ 1 show neutral selection (Z. Zhang, 2022). K_a/K_s ratios were calculated for *LUG* and *SEU* homologs of *A. thaliana*, *C. richardii*, *M. polymorpha*, and *P. patens* in relation to their respective homologs in the streptophyte algae *K. nitens*, thus spanning 800 million years of independent evolution (Bierenbroodspot *et al.*, 2024). The K_a/K_s ratios were calculated for both full-length genes, and for the individual identified protein domain regions (Fig. 6A – B). The strength of observed purifying selection was categorized according to Dong *et al.* (2019): Strong purifying selection: $K_a/K_s < 0.1$; moderate selection: $0.1 \leq K_a/K_s < 0.5$; weak negative selection: $0.5 \leq K_a/K_s < 0.9$.

Both *LUG* and *SEU* homologs showed a moderate purifying selection across their whole sequence (K_a/K_s ratio < 0.2 and < 0.4, respectively), and a generally strong purifying selection of $K_a/K_s \approx 0.1$ for the regions corresponding to functional protein domains. The intermediate region of *LUG* and the IDRs of *SEU* displayed comparatively higher K_a/K_s ratios with > 0.4 and ≈ 0.6, indicating a moderate or weak purifying selection, respectively. As a control of the K_a/K_s analysis, a second analysis comparing full length *LUG* and *SEU* sequences against the respective *C. richardii CrLUG* and *CrSEU1* homologs was performed, covering 450 million years of independent land plant evolution (Harris *et al.*, 2022). Similarly to the previous analysis, the *LUG* family exhibited a strong purifying selection ($K_a/K_s < 0.1$), while *SEU* showed a moderate negative selection of ≈ 0.3.

However, a closer look at the actual K_a and K_s values, and not only the calculated ratios, revealed that the synonymous substitution rates were quite high, ranging from 1.5 to 3.3 across all samples, with especially groups exhibiting low ratios possessing high K_s rates (Supplementary Table 1). Under high amounts of synonymous substitutions, a saturation of K_s values can occur, where the values can't be accurately calculated anymore, due to multiple substitutions covering each other. It has been suggested that this saturation only becomes noticeable at around $K_s = 2$ (Vanneste *et al.*, 2013), and causes underestimation of K_s values, with true rates of synonymous substitutions being higher than calculated (Gharib & Robinson-Rechavi, 2013; Vanneste *et al.*, 2013). Therefore, K_a/K_s ratios tend to be inflated under saturation, causing an elevated false positive rate when screening for genes under adaptive selection with $K_a/K_s > 1$ (Anisimova & Yang, 2007; Gharib & Robinson-Rechavi, 2013). Genes under neutral or purifying selection however have been suggested to be more robust against saturation effects, with little changes in the false positive rate under even high sequence divergence and thus amount of synonymous sub-

stitutions (Gharib & Robinson-Rechavi, 2013). Therefore, whilst the power of the K_a/K_s analysis is doubtlessly decreased under high K_s values, the general evidence of medium to strong purifying selection of *LUG* and *SEU* sequence across the green lineage should be supportable, even if only cautiously.

Given the relevance of the LUFS domain of LUG as a necessary and sufficient domain for SEU binding (Shrestha *et al.*, 2014; Sridhar *et al.*, 2004), and its absence as a detectable conserved domain in the Conserved Domain Database (J. Wang *et al.*, 2023), a consensus sequence of the LUFS domain was created. For this, the first 80 aa of 108 LUG homologs across streptophyte evolution were used as basis for the consensus (Fig. 7). As corroborated by the K_a/K_s analysis (Fig. 6A), large areas of the LUFS domain are strongly conserved across vast evolutionary distances, and remain largely invariant for a majority of analyzed LUG proteins. Especially noteworthy are two regions of strong sequence conservation: One is located directly at the N-terminus of the proteins from 6W to 19Y, while the other one ranges from 42P to 72S. Interestingly, the majority of the LisH domain, which is annotated to the residues 11 to 35, not only comprises a region with higher sequence variation than the surrounding LUFS region, but additionally shows considerable differences to the consensus of eukaryotic LisH domains (Gerlitz *et al.*, 2005). Several residues that are flexible in eukaryotes became more rigid in the plant LisH, such as L25 and F32, inferring information about the ancestral state of the LisH domain in the green plant lineage.

Taken together, both *LUG* and *SEU* genes and their proteins, and especially their functional domains like the LUFS region or the LDB domain, show considerable purifying selection and structural conservation for at least 800 million years of evolution, providing insight into the persistent selective pressure exerted upon both of these families throughout streptophyte algae and land plant evolution.

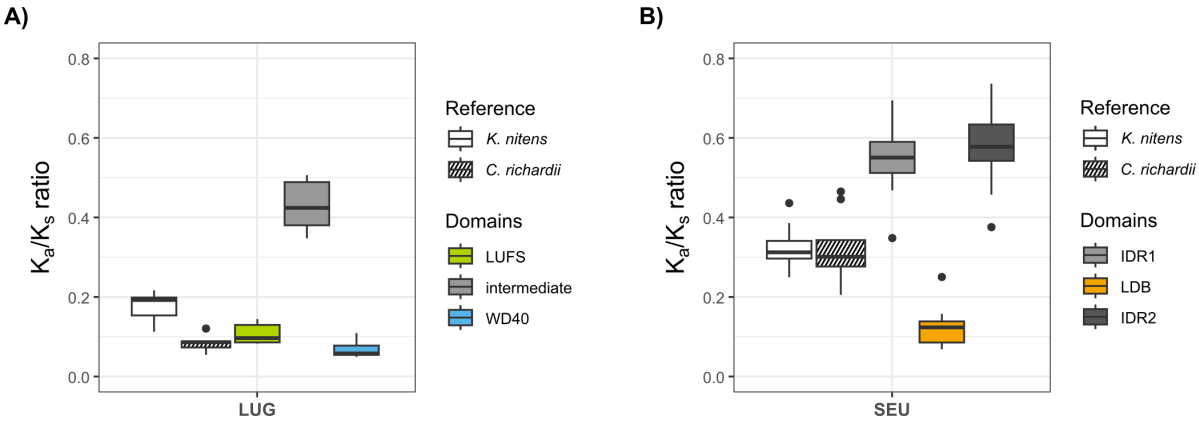


Figure 6: K_a/K_s analysis of LUG and SEU homologs across plant evolution. K_a/K_s analysis of **A)** full-length *LUG* sequences and individual protein domain regions, and of **B)** full-length *SEU* sequences and individual protein domain regions, in relation to *K. nitens* KnLUG and KnSEU, respectively. A comparative K_a/K_s analysis of whole *LUG* and *SEU* genes to *C. richardii CrLUG* or *CrSEU1*, respectively, was performed as well (striped plots).

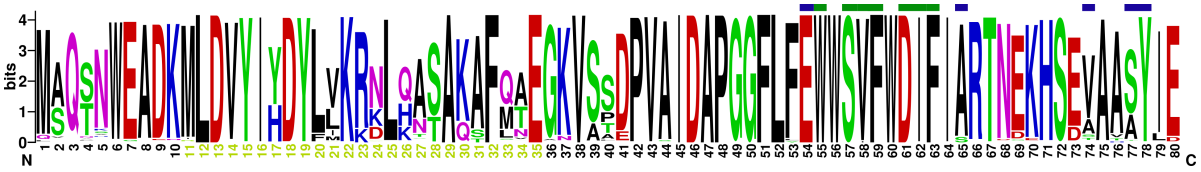


Figure 7: LUFS consensus sequence of 108 LUG homologs. Consensus sequence of the first 80 aa of the LUG N-terminus, showing predominant residues and residue frequencies at every position (Schneider & Stephens, 1990). The position of the LisH motif is colored lime. Bars on top signify consistently interacting residues in LUG – SEU complexes. Blue bars: Consistently interacting residues across one or two LUG homologs. Green bars: Consistently interacting residues across three to six LUG homologs.

4.3 The formation of the LUG – SEU module emerged at least 800 million years ago

So far, *in vivo* interaction between LUG and SEU homologs have been shown for the angiosperm species *A. thaliana* (Sitaraman *et al.*, 2008; Stahle *et al.*, 2009), *A. majus* (Navarro *et al.*, 2004), and rice (*Oriza sativa*; Yang *et al.*, 2019), but neither in a complete manner, testing the interaction between all present homologs with each other, nor for representatives outside of the angiosperms. Therefore, both *in silico* and *in vivo* analyses of the formation of the LUG – SEU module were performed across major land plant lineages, and for selected streptophyte algae lineages.

4.3.1 Digital gene expression analysis revealed a general, mostly uniform expression of both *LUG* and *SEU* homologs throughout major land plant lineages

As a first step, digital gene expression analysis was performed, to assess the general possibility of a formation of the LUG – SEU module in various tissues, and to see if differential expression between *LUG* or *SEU* homologs was present within the different species. For this, the expression levels of *LUG* and *SEU* homologs as transcripts per million (TPM) from published transcriptomes of *A. thaliana, C. richardii*, *M. polymorpha*, and *P. patens* (see section 3.2.6) were analyzed across both developmental stages, and in various different plant tissues (Fig. 8):

In *A. thaliana*, *AtLUH* was the overall strongest expressed *LUG* or *SEU* gene across all analyzed tissues, followed by *AtLUG* and *AtSLK2* (Fig. 8A). *AtSEU* and *AtSLK1* were expressed at a similar level to each other, while *AtSLK3* showed the lowest expression in comparison. Generally, both the *LUG* and the *SEU* families were tendentially stronger expressed in developing and mature flowers, and especially in the carpels, than in vegetative tissue, with a notably reduced expression in the embryo and stamen, and a nearly complete lack of gene expression in pollen. Interestingly however, analyzing the protein abundance in *A. thaliana* tissues (Mergner *et al.*, 2020) painted a slightly different picture: As expected from its expression levels, AtLUH was the most abundant LUG or SEU protein overall, while AtLUG and AtSEU showed slightly lower abundance levels (Fig. 8E). The three AtSLKs exhibited even lower protein levels, with especially AtSLK3 only being present in a

low quantity, at least in tissues were analysis was possible. While the exact scale range of iBAQ values depends on multiple different factors, the LUG and SEU proteins nevertheless showed an at least medium protein abundance level, with usually upwards of one million ($\log_2(20)$) copies per tissue (Mergner *et al.*, 2020; Zubarev, 2013). Interestingly, presence of low to medium amounts of all proteins could also be detected in stamen and pollen, indicating either a higher translation than surmised from the expression values of the genes alone, or some amount of protein movement between these tissues.

The fern *C. richardii* possessed a high, prominent expression of the *LUG* homolog *CrLUG*, while the *CrSEU* genes all showed a largely uniform expression across all tissues, with only *CrSEU2* being slightly more expressed than the others (Fig. 8B). Generally, no noticeable tissue specific expression patterns were present for either gene family.

While the *P. patens* gene *PpLUG1* showed the overall strongest gene expression, nearly no expression was recorded for its highly similar homolog *PpLUG2* (Fig. 8C). This could indicate a process of pseudogenization or gene silencing for *PpLUG2* after a duplication of the *PpLUG1/2* ancestor gene following a WGD event during the last 70 million years (Lang *et al.*, 2018), resulting in loss of expression, for example by accumulation of mutations in its promoter region due to relaxed selective pressure (Lynch & Conery, 2000; Soltis *et al.*, 2016). The *PpLUG3* and *PpLUG4* genes on the other hand were seemingly both maintained, with similar expression levels to each other, and to *PpSEU2*, while the other *SEU* homologs of *P. patens*, *PpSEU1*, *PpSEU3*, *PpSEU4*, *PpSEU5*, and *PpSEU6,* showed a comparatively slightly elevated expression overall. Similar to *C. richardii*, no general tissue specific *PpLUG* or *PpSEU* expression patterns were observable.

Finally, the expression of *MpLUG* in *M. polymorpha* was generally stronger than expression of *MpSEU1* and of *MpSEU2*; and expression of both families was elevated in developing, reproductive tissues, especially in antheridia and archegonia (Fig. 8D). Conversely, they were expressed at a lower level in vegetative tissues, and particularly in sperm cells.

Across land plants, both *LUG* and *SEU* homologs are usually expressed throughout the whole plant, with a sometimes elevated expression in reproductive tissues in comparison to vegetative ones. However, while *LUG* homologs show a tendentially stronger overall gene expression than *SEU* genes, no strong differential expression patterns between the *LUG* and the *SEU* families, nor within them, are noticeable beyond general global expression differences.

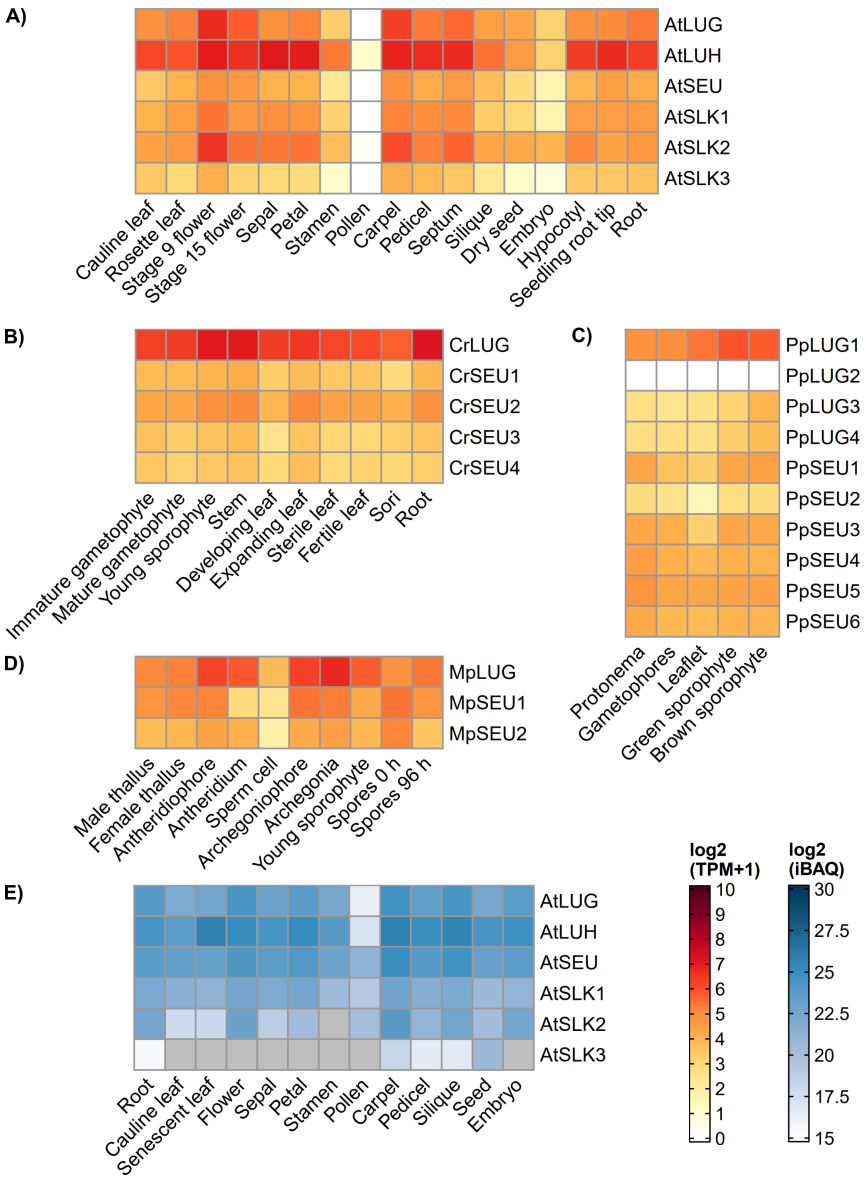


Figure 8: Comparative gene expression and proteome analysis of *LUG* and *SEU* homologs. A) – D) Gene expression of *LUG* and *SEU* homologs across different developmental tissues and stages, as $\log_2$(TPM + 1) values, and with identical color scales. **A)** Gene expression values for *A. thaliana* Col-0 (Mergner *et al.*, 2020); **B)** for *C. richardii* strain Hn-n (D. B. Marchant *et al.*, 2022); **C)** for *P. patens* Gransden 2004 (Lang *et al.*, 2018; Sreedasyam *et al.*, 2023); and for **D)** *M. polymorpha* 15 d male and female thallus (Melbourne strain) (Briginshaw *et al.*, 2022); Tak-1 sperm cells (Julca *et al.*, 2021), antheridiophores, and antheridia (Higo *et al.*, 2016); Tak-2 archegoniophores (Higo *et al.*, 2016) and archegonia (Hisanaga *et al.*, 2021); BC3 x Tak-1 13 d old sporophytes (Frank & Scanlon, 2015); and Cam1 x Cam2 spores (Bowman *et al.*, 2017). **E)** Protein abundance as $\log_2$(iBAQ) for *A. thaliana* Col-0 tissues (Mergner *et al.*, 2020).

4.3.2 *In silico* analysis of LUG – SEU complexes reveal conserved interacting protein regions and consistently interacting amino acid residues

Visualization and *in silico* analysis of LUG – SEU dimers across both major land plant and streptophyte algae lineages was performed, using AlphaFold modeling for all possible LUG and SEU homolog combinations of *A. thaliana, C. richardii*, the Charophyceae *Chara braunii*, *K. nitens*, and *M. polymorpha*, as well as for the *P. patens* homologs PpLUG1, PpLUG4, PpSEU1, PpSEU2, and PpSEU6 (Fig. 9; Supplementary Fig. 5 and 6; Abramson *et al.*, 2024). For dimer visualization, only the first 160 aa of the LUFS domains and interacting regions of the SEU homologs were pictured, as the LUFS region is necessary and sufficient for binding SEU (Shrestha *et al.*, 2014; Sridhar *et al.*, 2004), and the large stretches lacking defined structures present for both protein created a high amount of visual noise (Fig. 9A).

Generally, the interaction between the LUFS domains and SEU occurred mainly in the region of the LDB domain, and along the third and fourth α-helix of LUG (Fig. 9B – F; Supplementary Fig. 5). The third α-helix of LUG homologs usually had a lower positional confidence score than the other α-helices in this region (50 < pIDDT confidence ≤ 70), and thus was not included in Fig. 5A. It mainly interacted with short α-helices located near the central β-sheets of SEU, while the fourth α-helix of LUG was consistently significantly longer than the other LUG helices present, and interacted with one to two long α-helices of SEU, as well as sometimes with the LDB domain β-sheets. Interestingly, no interacting residues between SEU and the LisH domain were observed across multiple LUG homologs. However, as the LisH motif has been suggested to be a dimerization motif forming stable homodimers between two LisH domains (Gerlitz *et al.*, 2005; M. H. Kim *et al.*, 2004), a lack of interaction with SEU is not unexpected. Indeed, this analysis strongly suggests that the interaction between LUG and SEU proteins is mainly dependent on the conserved sequence region of the LUFS domain, rather than the LisH motif.

When mapping the position of interacting residues across the whole protein dimers, a conserved pattern can be observed (Fig. 9G; Supplementary Fig. 6): For LUG homologs, two areas of interaction are present across all analyzed dimers, namely the conserved region of the LUFS domain described above, and a secondary region of predicted interacting residues at the WD40 domains. For SEU homologs on the other hand, mapped interacting residues are located centrally within the protein, scattered in and around the LDB domain across a distance of approximately 250 aa. However, due to the general low positional confidence scores of IDRs and the cutoff values chosen for this analysis, interacting residues located within these IDRs can't be reliably detected using this method, and as such, further assessment of the reported necessity of the IDRs for the LUG – SEU module interaction requires different means (Sridhar *et al.*, 2004).

Mapping the position of LUFS-located residues that consistently interact with SEU across multiple LUG homologs, SEU interaction partners, and even tested species revealed that

a majority of these consistently interacting residues are located between 54E to 63F, and interact within at least three different LUG homologs (Fig. 7). As these residues are located within a region of the LUFS domain showing extreme sequence conservation across the green lineage, they are prime candidates for being critical residues in the LUG – SEU module formation.

In summary, *in silico* analysis of various LUG – SEU modules not only emphasized the general regions of LUG and SEU homologs crucial for dimer formation, but also revealed an important, conserved set of interacting residues within the LUFS domain that is vital for the binding of SEU homologs.

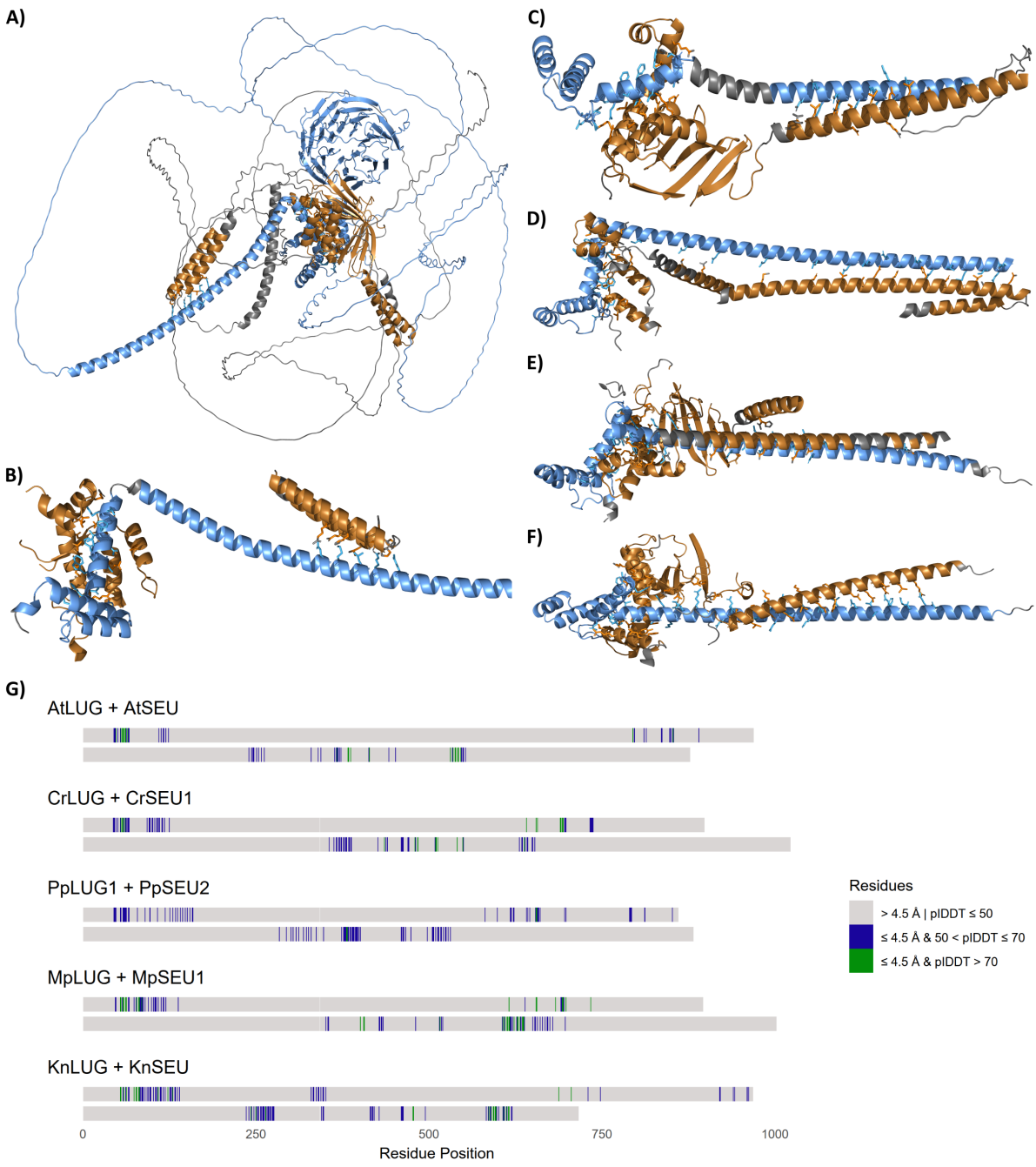


Figure 9: *In silico* prediction of LUG – SEU dimer interaction. Overview of interacting amino acid residues of LUG and SEU homologs within close proximity (≤ 4.5 Å) to partner chain. **A)** – **F)**: Modeled LUG (blue) – SEU (orange) protein complexes with visible interacting residues. Interacting residues have a pIDDT confidence of > 50 within 4.5 Å of a residue with a pIDDT confidence of > 50 of the partner protein chain (see Parvathy *et al.*, 2024). Grey regions have a pIDDT confidence < 50. **A)** Full length AtLUG – AtSEU dimer. **B)** – **F)**: First 160 N-terminal residues of LUG homologs and interacting SEU residues plus 20 surrounding atoms in both directions. **B)** AtLUG – AtSEU dimer; **C)** CrLUG – CrSEU1 dimer; **D)** PpLUG1 – PpSEU2 dimer; **E)** MpLUG – MpSEU1 dimer; **F)** KnLUG – KnSEU dimer. **G)** Position of interacting amino acid residues in LUG (upper) and SEU (lower) protein sequence. Grey residues are either not interacting, or have a pIDDT confidence of ≤ 50. Blue residues: Residues within 4.5 Å to partner chain, with pIDDT confidence between 50 and 70. Green residues: Residues within 4.5 Å to partner chain, with pIDDT confidence > 70.

4.3.3 Interactions between LUG and SEU homologs can be traced back to the Klebsormidiophyceae lineage of streptophyte algae

To systematically test homo- and heterodimerization between various LUG and SEU homologs of selected major green plant lineages *in vivo*, Y2H assays were performed (Fig. 10). All possible pairwise combinations within each species were tested on both normal triple dropout medium, and on dropout medium containing 2 mM of the competitive inhibitor 3-AT, which requires stronger protein interaction to allow growth of yeast cells. The only exception to this were the *P. patens* PpSEU3/4/5/6 proteins, as they were not present in the previous genome annotation v3.3 used as basis for identifying Y2H candidates (Lang *et al.*, 2018; see section 3.3.1).

A. thaliana possesses one LUG-, LUH-, and SEU-clade member each, as well as three SLK-clade proteins (Fig. 5). All six proteins showed extensive heterodimer formation in nearly all possible combinations and in both interaction directions, where possible (Fig. 10A). A lack of interaction was only observed for the combinations of AtLUH – AtSEU and for AtLUH – AtSLK3, the latter of which couldn't be tested due to autoactivation of both proteins. Most interactions between non-autoactivating proteins were visible in both combinations of activation domain (AD)- and binding domain (BD)-tagged proteins, and additionally, both AtLUG and AtSLK1 showed formation of homodimers. Assaying the interactions on more restrictive 3-AT containing medium abolished the autoactivation of AtLUH and AtSEU, and revealed strong protein interaction between those two proteins, as well as homodimer formation of AtSLK3 (Fig. 10B). Generally, most of the observed heterodimers, and the homodimer of AtSLK1, persistent under competitive yeast growth inhibition, indicating an overarching high strength of interaction for these protein dimers. Only six interactions were absent under these conditions, namely those between AtLUG – AtLUG, AtLUG – AtLUH, AtLUG –AtSLK2, AtSEU – AtSLK1, AtSLK1 – AtSLK2, and AtSLK2 – AtSLK3.

While many of these results validate previously published Y2H assays (Sitaraman *et al.*, 2008; Stahle *et al.*, 2009), there are differences present, too: Previous assays neither observed AtLUG – AtLUH interaction, nor AtLUG homodimers (Sitaraman *et al.*, 2008), or autoactivation of BD AtLUH constructs (Sitaraman *et al.*, 2008; Stahle *et al.*, 2009). Moreover, the absent interaction between AtLUH – AtSEU conflicts with previous reports as well (Sitaraman *et al.*, 2008; Stahle *et al.*, 2009). Some of these discrepancies can be explained by the differences in the employed assay media, as Sitaraman *et al.* (2008) used a quadruple dropout medium additionally lacking adenine, as well as an increased amount of 3-AT, thus requiring overall stronger interactions between the tested proteins to result in visible interaction. Unfortunately however, the exact medium composition used by Stahle *et al.* (2009) was not specified.

The singular LUG homolog of *C. richardii*, CrLUG, belonged to the LUG polytomy and interacted with both the SEU polytomy proteins CrSEU1, CrSEU2, and CrSEU3, as well as with

the SEU/SLK-sister-clade member CrSEU4 in a weak fashion (Fig. 10C). Additionally, it was the only tested *C. richardii* protein to show homodimerization, while CrSEU4 possessed weak autoactivation that was abolished under the influence of 3-AT. Contrastingly to *A. thaliana* or *M. polymorpha* SEU homologs, no homo- or heterodimerization between any CrSEU proteins was observed, and no protein interactions were present on 3-AT containing medium.

All four *P. patens* LUG homologs belong to the unresolved polytomy of bryophyte LUG proteins, and strikingly possessed only few interactions between each other, and with the SEU polytomy PpSEU1 and PpSEU2 proteins (Fig. 10D): Only PpLUG1, PpLUG4, and PpSEU2 showed any interactions at all, with all three proteins interacting with each other in all possible combinations, and both PpLUG4 and PpSEU2 additionally forming homodimers. In contrast, neither PpLUG2, nor PpLUG3 or PpSEU1 showed any dimerization at all. Further, all protein interactions were inhibited by the presence of 3-AT., similar to *C. richardii*.

The liverwort *M. polymorpha* possesses one bryophyte LUG protein and two SEU homologs, with MpSEU1 belonging to the SEU polytomy, while MpSEU2 is a member of the SEU/SLK-sister-clade. All three proteins formed heterodimers with each other, with MpLUG and MpSEU1 additionally forming homodimers (Fig. 10E). Due to autoactivation of MpSEU2, presence or absence of homodimer formation couldn't be elucidated. Although all heterodimers persisted on 3-AT containing medium, both homodimer formation and autoactivation were abolished under these conditions (Fig. 10F). The interactions between MpLUG with MpSEU1, and with MpSEU2, as well as the formation of MpLUG and MpSEU1 homodimers can also be corroborated by a Y2H assay performed by Q. Li (2022).

Finally, one algae LUG homolog and one algae SEU homolog are present in *K. nitens*. KnLUG and KnSEU interacted on both normal assay medium (Fig. 10G), and on medium containing 3-AT with each other (Fig. 10H). Autoactivation of KnLUG was abolished under presence of 3-AT, which additionally revealed the formation of a KnLUG homodimer.

To summarize, both the age and conservation of the LUG – SEU module across land plants and streptophyte algae could be demonstrated with an extensive degree of heterodimerization between LUG and SEU homologs in all representative species, spanning 800 million years of evolution in the green lineage (Bierenbroodspot *et al.*, 2024). This preservation of the interaction in combination with the observed purifying selection makes it a prime example of coevolution through strong reciprocal selective pressure, keeping both protein partners functional across vast time scales and ensuring the continued functionality of the LUG – SEU module (Lovell & Robertson, 2010). Additionally, extensive homodimerization could be observed for both families, which has not been described in previous interaction assays. Membership to specific phylogenetic clades however seemed independent from the ability of the homologs to form protein dimers.

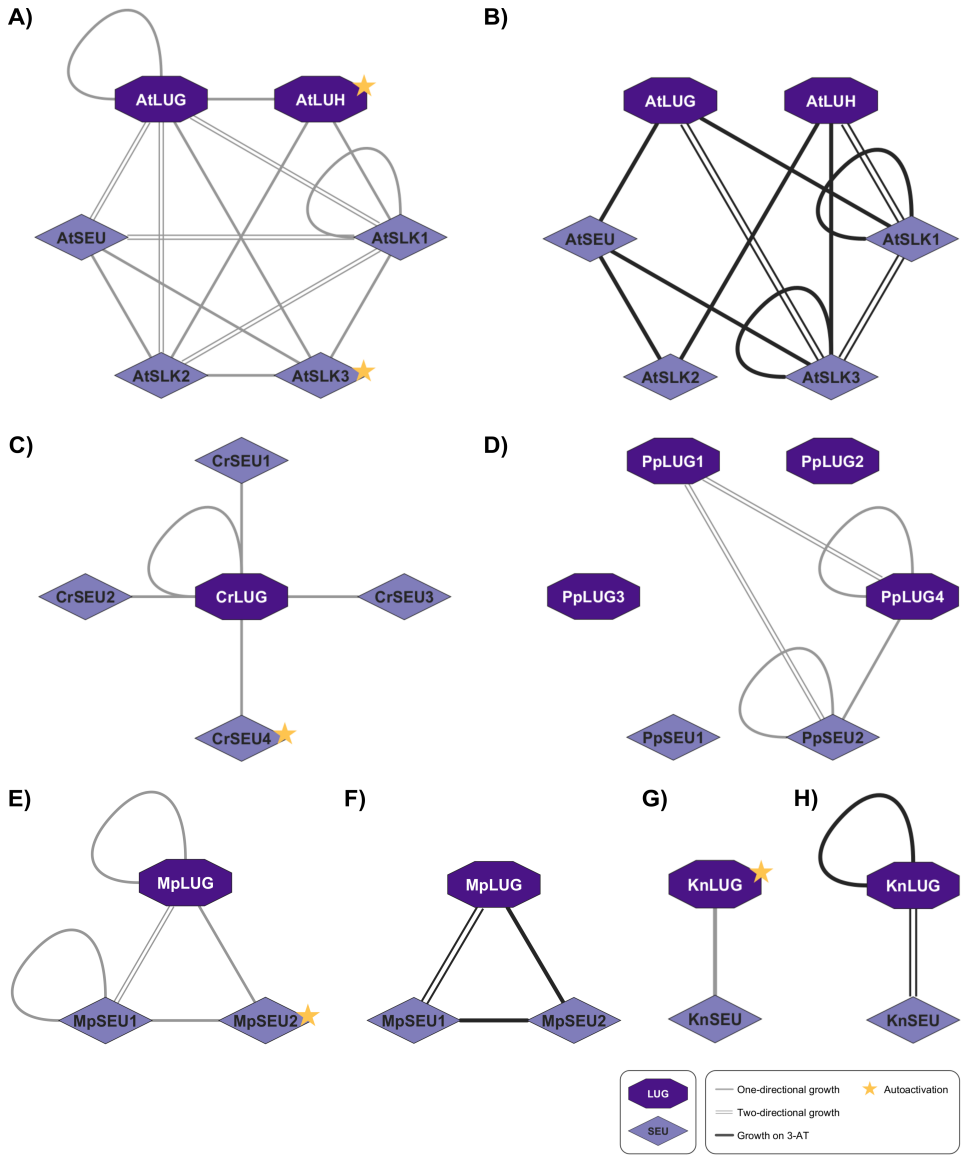


Figure 10: Interaction networks of LUG – SEU homologs across plant evolution. Networks of LUG and SEU homolog interactions analyzed in robot-assisted Y2H screening. Interactions for **A)** and **B)** *A. thaliana*; **C)** *C. richardii*; **D)** *P. patens*; **E)** and **F)** *M. polymorpha*; and **G)** and **H)** *K. nitens.* Interactions shown in **A)**, **C)**, **D)**, **E)**, and **G)** were observed on SD -LWH medium (grey lines), interactions in **B)**, **F)**, and **H)** on SD -LWH + 2 mM 3-AT (black lines). LUG and SEU homologs are distinguished by node color and shape, while edge type depicts interaction direction. One-directional growth: Growth of yeast colonies for one combination of AD protein X and BD protein Y, but not for AD protein Y – BD protein X. Two-directional growth: Growth of yeast colonies observed for both combinations of AD and BD plasmids. Yellow stars depict autoactivating BD constructs and were removed from the results. Autoactivation was only observed in on SD -LWH medium.

4.3.4 Independent BiFC assays verify several observed protein interactions for *C. richardii* and *M. polymorpha*

As Y2H assays are limited by their tendency to produce false positive and false negative results, both strict controls during the assay and independent verification of results are necessary to improve the quality of Y2H screens (Brückner *et al.*, 2009). Therefore, selected observed interactions were confirmed using a separated BiFC assay, and by referencing non-Y2H experiments in independent publications.

To support the Y2H results of the *C. richardii* LUG and SEU homologs, a BiFC assay of all possible combinations of CrLUG and CrSEU proteins with C-terminally fused YN or YC split-YFP tags was performed (Fig. 11). The negative controls and the full results of the assay can be found in Supplementary Fig. 7. Most observed fluorescence signals of reconstituted split-YFP following protein interaction were comparatively weak, especially in contrast to the fluorescence of the positive control, but BiFC assays are known to generally exhibit a wide range of fluorescence intensities between different protein partners (Kerppola, 2008).

Protein interactions were detected between CrLUG – CrSEU1, CrLUG – CrSEU2, and CrLUG – CrSEU3, verifying the results of the Y2H assay. Additionally, homodimerization of CrSEU2 and CrSEU4 proteins could be observed in the assay, both of which were not present in the yeast system. In case of CrSEU2, this might have been caused by a false negative result of the Y2H assay, while CrSEU4 homodimerization was not detectable due to autoactivation of CrSEU4 in yeast. Conversely, neither CrLUG homodimerization nor CrLUG – CrSEU4 interaction were observed in the BiFC assay, which might be caused by steric hindrance of the YC and YN tags, interfering in the reconstitution of the split-YFP protein (Kerppola, 2008).

The interaction between *M. polymorpha* MpLUG and MpSEU1 was previously verified by Q. Li (2022), via both a BiFC assay, and by Co-Immunoprecipitation (CoIP) in *N. benthamiana* leaves.

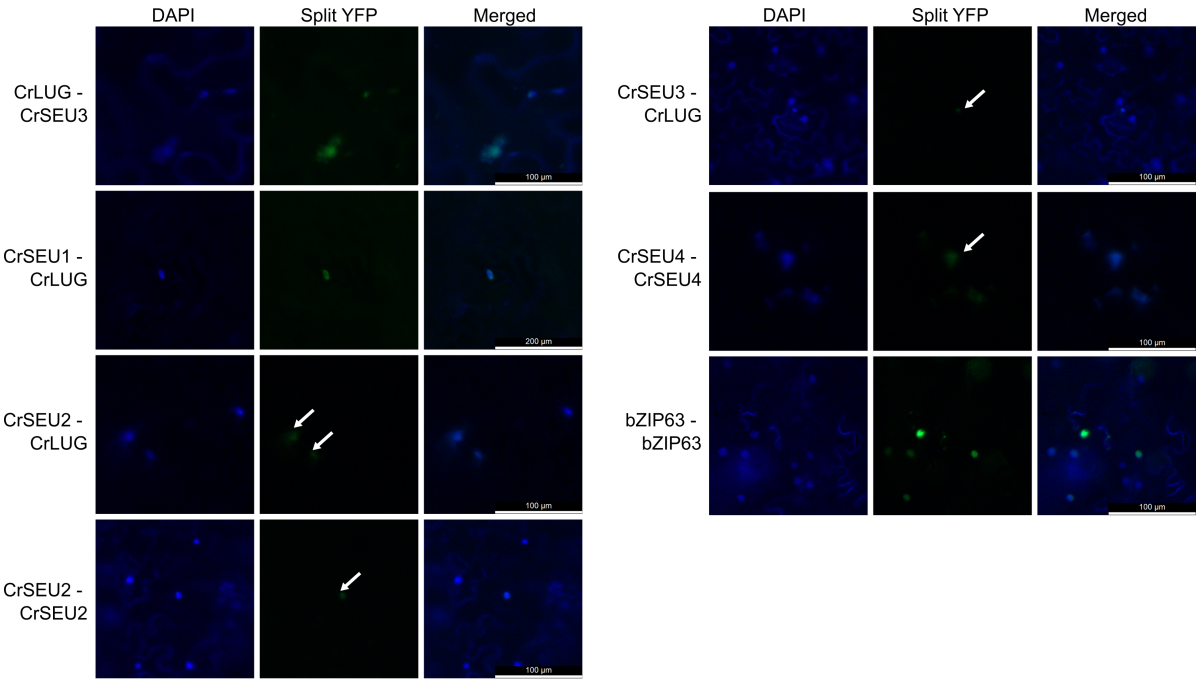


Figure 11: Independent BiFC assay of *C. richardii* CrLUG and CrSEU interactions. Selected results of BiFC assays of *C. richardii* CrLUG and CrSEU homolog interactions in *N. benthamiana*, with row-wise protein combinations as YC-tagged protein X – YN-tagged protein Y. „DAPI“ column: Fluorescence signal of nuclei stained with 100 $\frac{\mu g}{ml}$ DAPI. „Split-YFP“ column: Split-YFP protein fluorescence. „Merged“: Overlay between DAPI and split-YFP fluorescences, with row-wise scale bars. White arrows depict difficult to see YFP fluorescence signals. bZIP63 was used as a positive control, while negative controls are shown in Supplementary Fig. 7.

4.4 Heterodimerization between LUG homologs and MADS-box proteins predates the emergence of land plants

Given the presence and conservation of the LUG – SEU module throughout the green lineage, the question arises if their interaction with other known interaction partners was a conserved function since at least the Klebsormidiophyceae, or if such interactions emerged later during the evolution of land plants. Therefore, Y2H assays partnering LUG and SEU homologs with MIKC type MADS-box proteins were performed, as especially MIKC^{C} types like AP1 and SEP3 are known to be critical interaction partners of the LUG – SEU module in *A. thaliana* flower development (Conner & Liu, 2000; Franks *et al.*, 2002; Sridhar *et al.*, 2006). As, up until recently, the MIKC groups were considered to be more closely related to each other than to the M-type (Gramzow & Theissen, 2010; Qiu & Köhler, 2026), the interaction between LUG, SEU, and MIKC* proteins across streptophytes were analyzed as well, especially since their interactions in *A. thaliana* remain largely unexplored.

The MADS-box proteins for this assay were chosen either to allow comparison with already published interactions, especially with independent Y2H assays in *A. thaliana*, or picked randomly from a list of published MADS-box proteins. Both MIKC^{C} and MIKC* types were chosen, and for *P. patens*, two M-type MADS-box proteins, PPTIM2 and PPTIM3, were selected too, due to their notably strong, localized expression in sporophytes (Supplementary Fig. 1D).

In *A. thaliana*, both protein and genetic interactions between AtLUG and AtSEU with different MIKC^{C}s are well studied, however, comparatively little is known about protein interactions for AtLUH and the AtSLKs, and with the MIKC* group (Di Marzo *et al.*, 2022; Gregis *et al.*, 2006; Herrera-Ubaldo *et al.*, 2023; Sridhar *et al.*, 2006). Most of the tested MADS-box proteins showed interactions with multiple LUG or SEU homologs (Fig. 12A), except for the MIKC^{C} group protein AtSTK, which failed to interact with any tested LUG or SEU protein, as previously reported (Di Marzo *et al.*, 2022; Herrera-Ubaldo *et al.*, 2023). The MIKC^{C} type AtSEP2 exhibited heterodimer formation with all tested proteins except with AtLUG and AtSLK2, and similarly, AtSEP3 interacted with all LUG and SEU homologs besides AtLUH. AtAG however only interacted with the SLK proteins AtSLK1 and AtSLK3, but neither with AtLUG nor with AtSEU, as previously established (Conner & Liu, 2000). Both of the MIKC* type proteins AGAMOUS-LIKE 66 (AtAGL66) and AtAGL104 possessed identical interaction patterns, forming protein heterodimers with all tested LUG and SEU homologs. Due to autoactivation of all involved proteins, interactions between the MIKC* group members and AtLUH or AtSLK3 couldn't be elucidated. When performing the Y2H assay on 3-AT containing medium, only the dimer formation between AtSLK1 with either AtAGL66 or AtAGL104 persisted, indicating a stronger binding between those proteins than present in any other LUG / SEU – MADS-box dimer tested. Additionally, the MADS-box TFs AtSEP2, AtAG, and AtAGL104 showed yeast growth when paired with an empty

BD vector, indicating capacity for false-positive results. While several tested MADS-box proteins possessed this autoactivation of yeast growth when paired with an empty BD vector, it was not present for all possible protein partners in the assay, placing it in stark contrast to autoactivation with empty AD vectors. The latter consistently caused yeast growth across all replicates regardless of the tested partner protein, thus rendering the results unusable, while empty BD autoactivators didn't exhibit this pervasive spurious yeast growth. Therefore, it was decided to not outright remove them from the analysis across the board, but instead to mark such empty BD autoactivators accordingly, to keep it in mind during the analysis of their results.

Both MIKC^C and MIKC* proteins of *C. richardii* formed heterodimers with C-fern SEU homologs, while binding to CrLUG was generally absent, with the exception of the MIKC* group member CRM13 (Fig. 12B). Overall, only a low number of interactions were visible, most of which were between the SEU/SLK-sister protein CrSEU4 and both types of MADS-box proteins. Only the MIKC^C type protein CMADS1 and the MIKC* member CRM16 lacked heterodimer formation with CrSEU4, with the former showing no protein interactions at all, and the latter interacting with CrSEU1 and CrSEU3 instead. Additionally, an interaction between CrSEU3 and CRM13 was present, too. Interestingly, no heterodimers could be recorded between any MADS-box proteins and CrSEU2. Both CRM1 and CRM13 showed empty BD autoactivation, and none of the tested protein dimers persisted on 3-AT containing medium.

Nearly all observed interactions for *P. patens* MADS-box proteins occurred with PpLUG4 and with PpSEU2 (Fig. 12C): The MIKC^C group members PPM1 and PPM2 interacted with both of those proteins, while PPMADS1 only interacted with PpSEU2. Additionally, both PPM1 and PPMADS1 showed autoactivation with empty BD vectors. For the MIKC* proteins, PPM3 and PPM9 formed heterodimers with PpLUG4 and with PpSEU2, while PPM3 further interacted with PpLUG1 as well. Interestingly, PPMADS2 showed protein interaction with PpLUG3 and PpSEU2 instead, as well as autoactivation in BD position. Finally, the M-type MADS-box protein PPTIM2 didn't exhibit any protein dimer formation, while PPTIM3 was an autoactivator in BD position and interacted with PpLUG1, PpLUG4, PpSEU1, and PpSEU2. None of the described interactions were present on the more stringent medium containing 3-AT.

M. polymorpha encodes for only two MADS-box proteins, the MIKC^C protein MpMADS2 and the MIKC* protein MpMADS1. While the former only interacted weakly with MpSEU1, extensive heterodimer formation with all tested LUG and SEU homologs could be shown for MpMADS1, which also persisted on 3-AT containing medium, suggesting a certain strength of the formed protein dimers (Fig. 12D). While MpMADS1 possessed autoactivation when paired with an empty AD vector, it was abolished under the influence of 3-AT.

The streptophyte algae *K. nitens* possesses only a single MIKC type protein predating the divergences in the lineage leading to land plants (Kaufmann *et al.*, 2005; Kwantes *et al.*,

2012; Qiu & Köhler, 2026). This KnMADS protein exhibited weak heterodimer formation with KnLUG, but interestingly not with KnSEU (Fig. 12E).

In summary, interactions between LUG homologs and MIKC type MADS-box proteins were present across the streptophyte lineage, while interactions with the SEU family can be traced back to at least the MRCA of land plants. The vast majority of tested MADS-box proteins interacted with at least one LUG or SEU homolog, and oftentimes with both families. Likewise, nearly all LUG and SEU homologs interacted with at least one MADS-box proteins, without any detectable preference for either the $MIKC^C$ or the MIKC* family. As seen in their gene expression, the expression patterns between *LUG*, *SEU*, and MADS-box proteins generally overlap in most tissues (Fig. 8, Supplementary Fig. 1), indicating a stable, albeit weak, interaction between these families across hundreds of million of years of plant evolution, and potentially across most tissues and developmental stages.

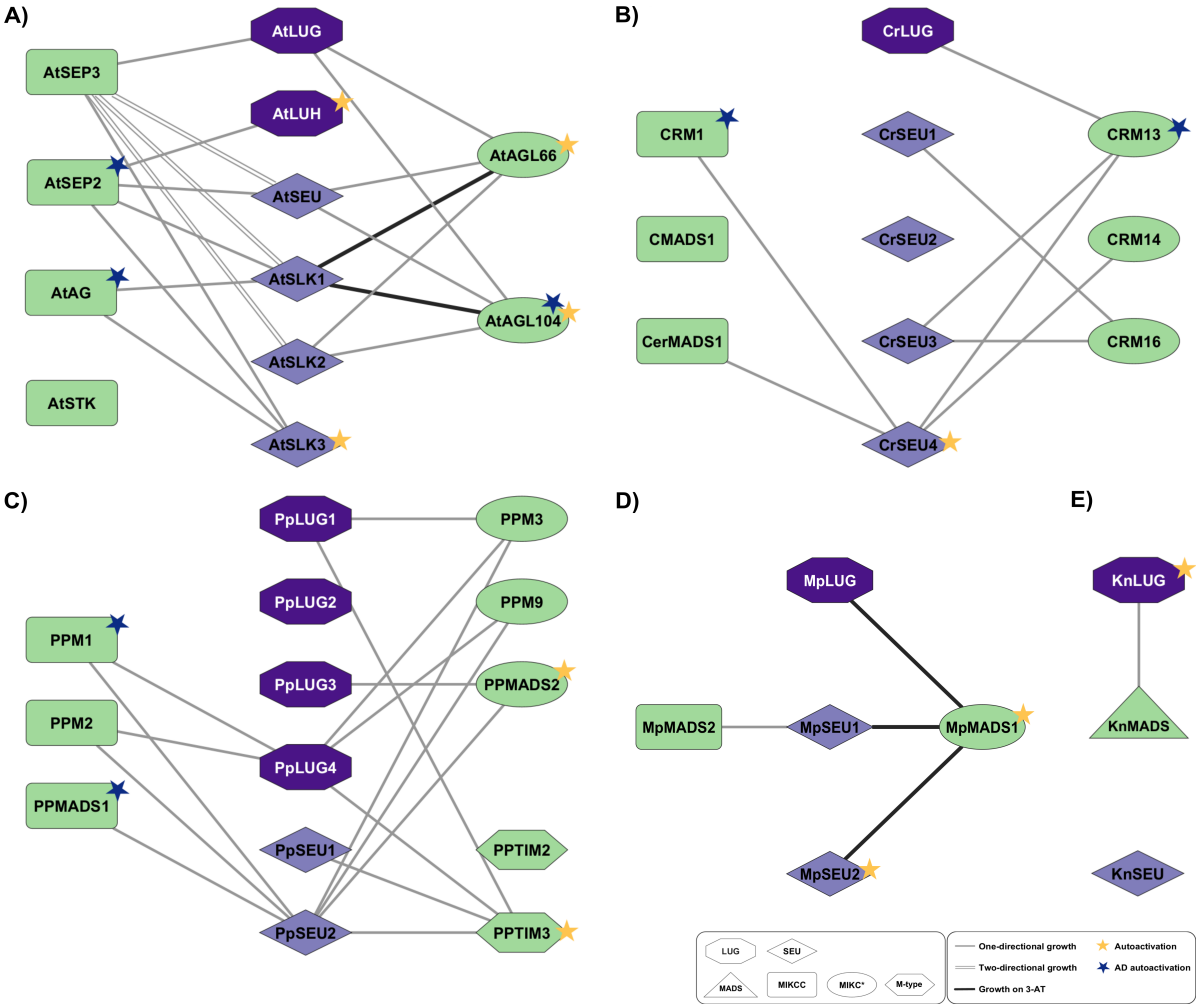


Figure 12: LUG / SEU – MADS-box protein interaction networks. Protein interaction networks of LUG / SEU proteins with MADS-box TFs. Interactions were analyzed in robot-assisted Y2H screening on SD -LWH (grey lines) and SD -LWH + 2 mM 3-AT (black lines) media. Interactions for **A)** *A. thaliana*; **B)** *C. richardii*; **C)** *P. patens*; **D)** *M. polymorpha*; and **E)** *K. nitens*. LUG and SEU homologs and MADS-box types are distinguished by node color and shape, while edge type depicts interaction direction. One-directional growth: Growth of yeast colonies for one combination of AD protein X and BD protein Y, but not for AD protein Y – BD protein X. Two-directional growth: Growth of yeast colonies observed for both combinations of AD and BD plasmids. Yellow stars depict autoactivating BD constructs and were removed from the results. Blue stars signify autoactivation of the AD construct. Autoactivation was only observed in on SD -LWH medium.

4.5 The *PpLUG3* and *PpLUG4* genes of *P. patens* function together with auxin in crucial protonema developmental transitions

To elucidate the functions of LUG and SEU proteins *in planta*, CRISPR/Cas9-mediated mutagenesis of *LUG* and *SEU* homologs of *P. patens* was performed. Due to time constraints, only mutants of the *pplug3 pplug4* double target CRISPR line could be analyzed; as such, mutant lines of the other *PpLUG* and *PpSEU* genes, as well as higher order mutant lines can't be covered here (see section

4.5.1 Analyzed *P. patens pplug3 pplug4* mutant lines encompass both short deletions and full protein knockouts

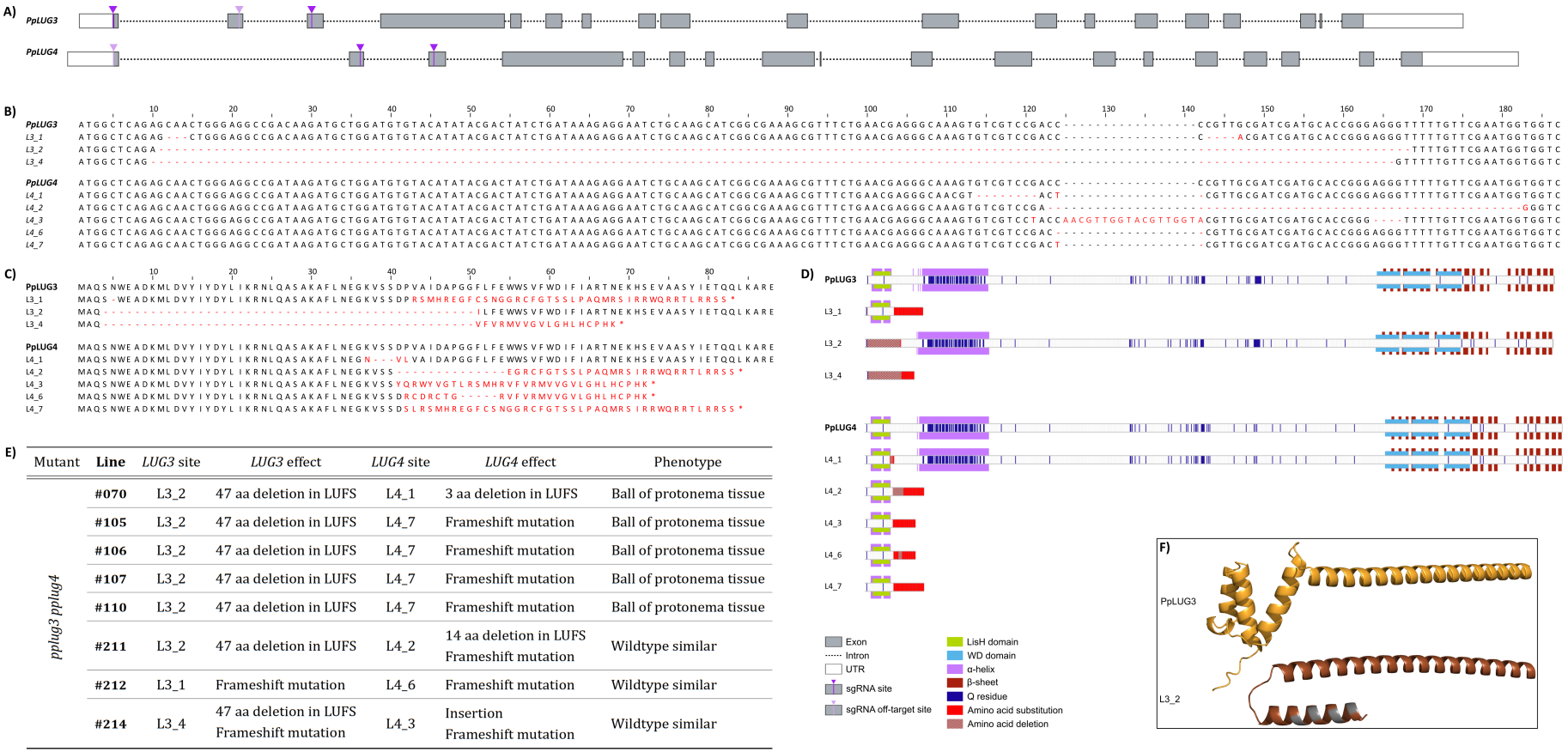
Both *PpLUG3* and *PpLUG4* genes were targeted by two sgRNAs each, whose complementary sequences were located within the first three exons of the respective gene. Due to the high sequence similarity of the *PpLUG3* and *PpLUG4* CDSs (96.7 % similarity, 2.1 % gaps in alignment; Madeira *et al.*, 2024; Smith & Waterman, 1981), their sgRNAs also possessed capacity for off-target effects in the respective other gene, with the sgRNA1 target of *PpLUG3* also affecting *PpLUG4*, and vice versa (Fig. 13A).

In total, eight individual *pplug3 pplug4* mutant lines were analyzed, which can be split into five different genotype groups regarding their effects on the PpLUG3 and PpLUG4 proteins (Fig. 13E):

1) Line #070: Loss of a region of the PpLUG3 LUFS domain and six mutated aa in PpLUG4, including a 3 aa deletion.
2) Lines #105, #106, #107, and #110: Loss of a region of the PpLUG3 LUFS domain and knockout of PpLUG4 due to a frameshift mutation 40 aa upstream of the N-terminus.
3) Line #211: Loss of a region of the PpLUG3 LUFS domain paired with a frameshift mutation of PpLUG4, similar to but differing from the previous group.
4) Line #212: Frameshift mutations of both PpLUG3 and PpLUG4 after around 40 aa, leading to a knockout of both genes.
5) Line #214: Deletion and frameshift mutation in *PpLUG3* at the start of the gene, resulting in a complete loss of PpLUG3, plus an insertion in the second exon of *PpLUG4*, leading to a frameshift mutation.

Each individual mutant allele was assigned an ID (Fig. 13B – C); and Fig. 13D contains an overview over the effects of the individual mutations on the overall protein structure and especially the conserved functional domains of PpLUG3 and PpLUG4. Briefly, three different mutant alleles for *PpLUG3* and five for *PpLUG4* were observed: The *PpLUG3* mutations either caused a frameshift and subsequent loss of PpLUG3 in the region following the LisH motif (ID L3_1; group **4)**), a loss of the LisH domain and of the surrounding region, but an otherwise intact PpLUG3 protein (ID L3_2; groups **1)**, **2)** and **3)**, Fig. 13F),

or a complete loss of PpLUG3 beyond the first four aa (ID L3_4; group **5)**). For *PpLUG4*, mutations either caused a short, six aa long mutation in the LUFS domain encompassing a three aa long deletion (ID L4_1; group **1)**), or a frameshift mutation directly after the LisH motif. For this second mutation, four different variations were observed (IDs L4_2, L4_3, L4_6, L4_7; groups **3)**, **5)**, **4)**, and **2)**, respectively). Fig. 13F depicts the first 130 aa of WT PpLUG3 in comparison to the protein of the L3_2 deletion, exhibiting a notable loss of N-terminal α-helices.



E)

Mutant	Line	*LUG3* site	*LUG3* effect	*LUG4* site	*LUG4* effect	Phenotype
pplug3 pplug4	**#070**	L3_2	47 aa deletion in LUFS	L4_1	3 aa deletion in LUFS	Ball of protonema tissue
	#105	L3_2	47 aa deletion in LUFS	L4_7	Frameshift mutation	Ball of protonema tissue
	#106	L3_2	47 aa deletion in LUFS	L4_7	Frameshift mutation	Ball of protonema tissue
	#107	L3_2	47 aa deletion in LUFS	L4_7	Frameshift mutation	Ball of protonema tissue
	#110	L3_2	47 aa deletion in LUFS	L4_7	Frameshift mutation	Ball of protonema tissue
	#211	L3_2	47 aa deletion in LUFS	L4_2	14 aa deletion in LUFS Frameshift mutation	Wildtype similar
	#212	L3_1	Frameshift mutation	L4_6	Frameshift mutation	Wildtype similar
	#214	L3_4	47 aa deletion in LUFS Frameshift mutation	L4_3	Insertion Frameshift mutation	Wildtype similar

Figure 13: Position, structure, and effects of *pplug3* and *pplug4* mutant alleles. A) Whole *PpLUG3* and *PpLUG4* genes with exon (grey box), intron (dashed line), and untranslated region (UTR, white box) positions. Purple arrows indicate sgRNA target sites, while mauve arrows show potential off-target sites of sgRNAs belonging to the respective other gene. **B)** Alignment of the first ~180 bp of *PpLUG3* and *PpLUG4* CDSs, and of observed mutations of these sequences. Each mutation was assigned an unique ID, and differences to WT CDSs are marked in red. **C)** Alignment of the first 88 aa of PpLUG3 and PpLUG4, and of translated mutant sequences. Differences in aa sequence between WT and mutants were marked in red. **D)** Visualization of the effects of the recorded mutant alleles on the protein sequences, with marked Q residues (dark blue lines), secondary protein structures at a pIDDT confidence ≥ 70 (purple = α-helices; dark red = β-sheets), and conserved protein domains (green = LisH domain, light blue = WD40 domains). Regions with aa substitutions were marked in red, while aa deletions were marked with red and grey stripes. **E)** Table of all eight mutant lines used, listing the respective mutant allele IDs and brief descriptions for both *PpLUG3* and *PpLUG4* loci, as well as a simplified plant habit phenotype. **F)** AlphaFold model of the first 130 aa of PpLUG3 (orange) and of the L3_2 deletion (brown), with residues below a pIDDT confidence of 50 colored grey (Abramson *et al.*, 2024).

4.5.2 Vegetative and reproductive phenotyping of *P. patens* mutants revealed two principal *pplug3 pplug4* phenotypes

The eight *pplug3 pplug4* mutant lines exhibited two main phenotypes (Fig. 13E and 14A – C): The mutant plants of groups **1)** and **2)** that lack the LisH domain of PpLUG3 and show either a disruption of the PpLUG4 LUFS domain, or absence of PpLUG4, were small, densely packed clumps of protonema-similar cells, and therefore dubbed „balls of protonema tissue", or „ball" phenotypes, for short. Those mutants belonging to groups **3)**, **4)**, and **5)** however, which lacked PpLUG4 and either only the PpLUG3 LisH domain, or the whole PpLUG3 protein, appeared outwardly similar to the *P. patens* wildtype and were consequently named „wildtype similar" or „WT similar". To analyze the exact effects of the *pplug3 pplug4* mutations on the development of *P. patens* in more detail, and to elucidate if there were more nuanced differences within these two principal phenotypes present, in depth phenotyping of protonema and gametophore growth parameters and of the reproductive cycle was performed.

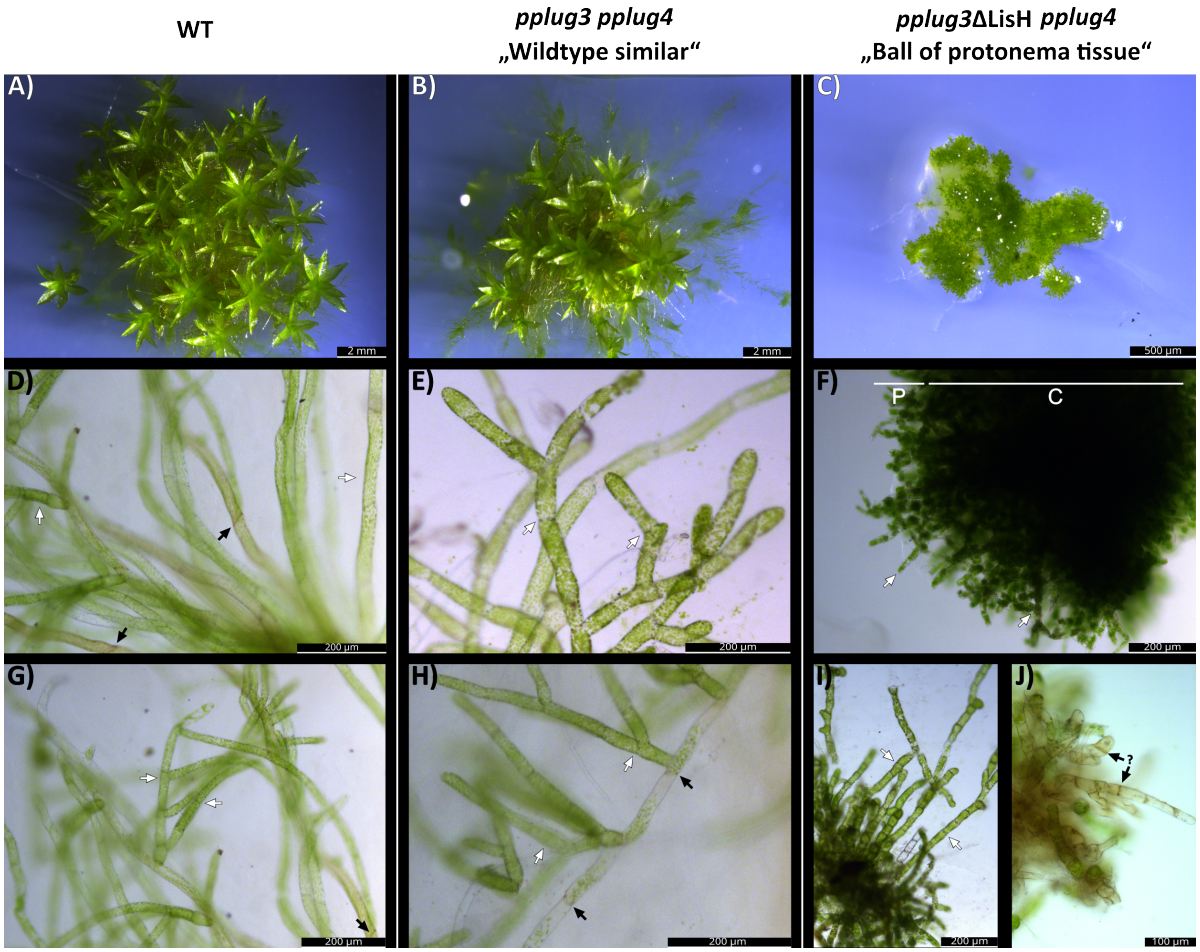


Figure 14: Phenotypic analysis of vegetative tissues of the two principal *pplug3 pplug4 P. patens* mutant phenotypes. Left column (**A)**; **D)**; **G)**): WT plants. Middle column (**B)**; **E)**; **H)**): *pplug3 pplug4* mutants showing a „WT similar" phenotype. Right column (**C)**; **F)**; **I)**; **J)**): *pplug3*ΔLUFS *pplug4* mutants resembling small „balls of protonema tissue". Depicted are **A) – C)** the whole gametophyte, **D) – I)** chloronema (white arrows) and caulonema (black arrows) cells, and **F)** the two zones of the „ball" phenotypes: (C)enter and (P)eriphery. **J)** Protonema cells with unclear identity.

4.5.2.1 PpLUG3 and PpLUG4 govern protonema development, transition, and gametophore initiation in *P. patens*

As a first step, the area covered by the gametophyte after four weeks of growth was measured, to obtain a parameter for gametophyte spreading. The coverage of three individual gametophytes for each line was determined and compared against the size of the area covered by the WT gametophytes, revealing that all „ball of protonema tissue“ mutant lines exhibited a significantly reduced gametophyte growth, spreading across only $\sim \frac{1}{50}$ of the area covered by WT gametophytes (Fig. 14A – B). Thus, these „ball“ phenotype lines exhibited a strongly diminished ability to colonize their surroundings. In contrast, „WT similar“ mutant lines showed no significant differences in gametophyte growth rates compared to the WT gametophyte.

As protonema are the first emerging tissue during *P. patens* development, and among other things responsible for colonization and foraging (Jaeger & Moody, 2021; Rensing *et al.*, 2020), the growth and developmental characteristics of protonema for the different mutant lines were analyzed: Generally, the protonema of the „WT similar“ mutant lines closely resembled the protonema of the WT, while those of the „balls of protonema tissue“ noticeably differed from WT protonema cells and from their behaviour (Fig. 15C). For one, the protonema of WT and „WT similar“ gametophytes grew freely through the medium in a primarily two dimensional (2D) fashion, with no protonema growth above the culture medium or straight downwards (Fig. 14D – E, G – H). The protonema of „ball“ phenotypes however grew in a pronounced three dimensional (3D) manner, forming dense „balls“ of clumped cells, and even growing outside of the medium (Fig. 14C). These „balls“ can be divided into two sections, namely the tightly packed center of the clump, containing the majority of protonema cells, and the periphery of the gametophyte (Fig. 14F). From the periphery, cells were able to grow outwards in all directions, but only for a short distance. However, during prolonged growth, and especially under short day conditions, outgrowth of longer chains of protonema away from the periphery could be observed (Fig. 14I).

There are multiple types of protonema tissue, with especially chloronema and caulonema playing major roles during the early stages of the *P. patens* life cycle. They differ in multiple characteristics, like cell wall orientation, chloroplast content, or function (Jaeger & Moody, 2021; Rensing *et al.*, 2020). For all mutant lines and the WT, the presence of protonema cells with transverse cell plates and a high chloroplast content was observed, indicating that these cells were most likely chloronema (Fig. 14D – I and 15C). However, the chloronema cells of both principal phenotypes are not entirely identical to WT chloronema, with especially the cell sizes of „ball“ phenotype lines diverging significantly from the WT (Fig. 15D – E): With only $\frac{1}{2}$ to $\frac{2}{3}$ of the WT length, their chloronema cells were consistently significantly shorter, as well as significantly thicker by a few µm than WT chloronema. Similarly, while the chloronema of „WT similar“ lines were not always

significantly different in both length and width to the WT, those that exhibited differences tended to be slightly longer and wider than WT chloronema.

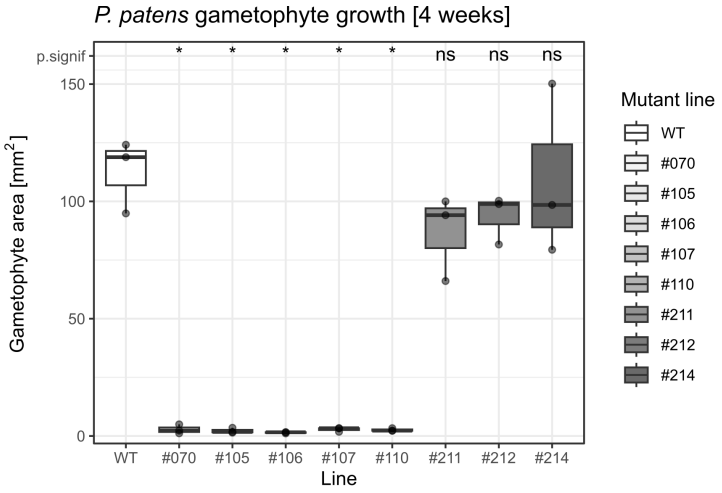
Caulonema, the second main type of protonema, were identified in WT and „WT similar“ lines due to their oblique cell plates and low chloroplast numbers (Fig. 15C; Jaeger & Moody, 2021; Rensing *et al.*, 2020). Additionally, for both groups, rhizoids were present as well, which emerge from the base of the gametophores and are similar to caulonema, but contain immature plastids and show a brown pigmentation (Kofuji & Hasebe, 2014). For the „ball of protonema tissue“ phenotypes however, neither caulonema nor rhizoids could be unequivocally identified. While there were cells present resembling caulonema, especially after eight weeks of growth, they exhibited critical deviations from the normal caulonema characteristics (Fig. 14J): Many of the potential caulonema cells contained only few chloroplasts, but also a brownish coloration and transverse cell plates, instead of oblique ones. Therefore, the exact identity of these cells remains ambiguous, which together with the prevalence of chloronema cells indicates a severe defect in caulonema identity establishment for these mutant lines.

Another major difference between the two *pplug3 pplug4* mutants to the *P. patens* WT was the presence and initiation of gametophores. None of the studied „ball of protonema tissue“ lines exhibited the formation of any structures resembling either gametophores, stems, or leaflets at all (Fig. 14C and 15F). „WT similar“ mutants on the other hand exhibited a noticeable delay in gametophore initiation from protonema: While the whole gametophyte area was covered by mature gametophores for the WT (Fig. 14A), only the inner region of the gametophyte was covered for „WT similar“ lines, leaving an outer „ring“ of spread-out protonema, but without visible gametophore growth (Fig. 14B). Additionally, many of the present gametophores were immature and still developing. Full coverage of the gametophyte with mature gametophores was only achieved after eight weeks of growth, and no further differences between „WT similar“ mutant and the WT regarding the overall structure or specific organs of the gametophores themselves were present (Fig. 15F).

In summary, mutations of *PpLUG3* and *PpLUG4* affect chloronema growth, protonema identity specification, and the initiation of gametophores from protonema, with especially the „ball“ phenotype negatively influencing the ability of the gametophyte to colonize their surroundings, the establishment of caulonema identity, and abolishing gametophore initiation altogether.

Line	Simplified genotype		Habit
	PpLUG3	*PpLUG4*	
WT	—	—	Wildtype
#070	*pplug3*ΔLUFS	*pplug4*Δ3aa	Ball of protonema tissue
#105	*pplug3*ΔLUFS	*pplug4*	Ball of protonema tissue
#106	*pplug3*ΔLUFS	*pplug4*	Ball of protonema tissue
#107	*pplug3*ΔLUFS	*pplug4*	Ball of protonema tissue
#110	*pplug3*ΔLUFS	*pplug4*	Ball of protonema tissue
#211	*pplug3*ΔLUFS	*pplug4*	Wildtype similar
#212	*pplug3*	*pplug4*	Wildtype similar
#214	*pplug3*	*pplug4*	Wildtype similar

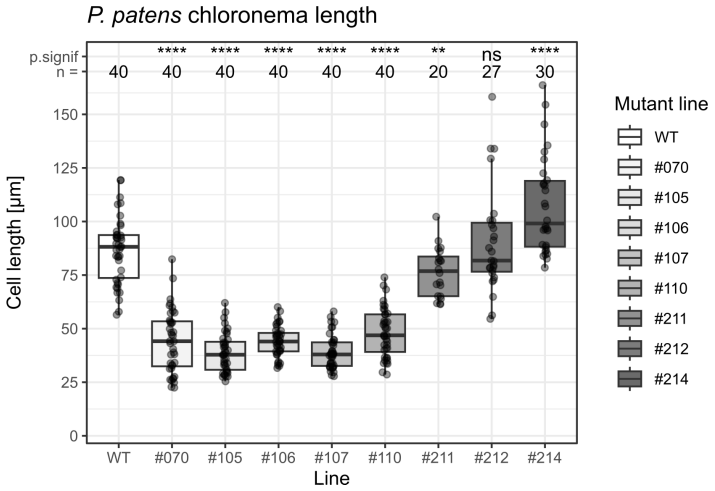
A) Simplified genotype and plant habit of *pplug3 pplug4 P. patens* mutants.



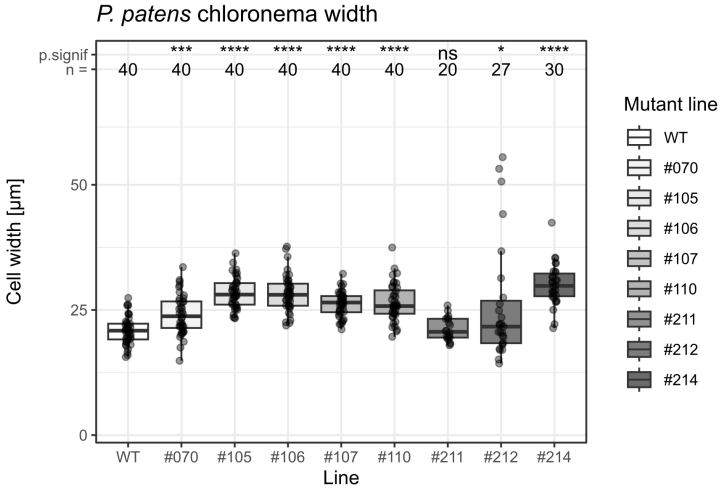
B) Gametophyte spread after four weeks of growth

Line	Presence	Growth dimension	Growth pattern			Cell types	
			4 weeks	8 weeks	Short day	4 weeks	8 weeks
WT	Present	2D growth	Free	Free	Free	Chl.; Cau.; Rhi.	Chl.; Cau.; Rhi.
#070	Present	3D growth	Clumped	Clumped	Outgrowing	Chl.*	Chl.*; Cau.*
#105	Present	3D growth	Clumped	Clumped – Outgrowing	Outgrowing	Chl.*; Cau.*	Chl.*; Cau.*
#106	Present	3D growth	Clumped	Clumped – Outgrowing	Outgrowing	Chl.*	Chl.*; Cau.*
#107	Present	3D growth	Clumped	Clumped	Outgrowing	Chl.*	Chl.*; Cau.*
#110	Present	3D growth	Clumped	Clumped – Outgrowing	Outgrowing	Chl.*	Chl.*; Cau.*
#211	Present	2D growth	Free	Free	Free	Chl.*; Cau.; Rhi.	Chl.*; Cau.; Rhi.
#212	Present	2D growth	Free	Free	Free	Chl.*; Cau.; Rhi.	Chl.*; Cau.; Rhi.
#214	Present	2D growth	Free	Free	Free	Chl.*; Cau.; Rhi.	Chl.*; Cau.; Rhi.

C) Protonema growth characteristics of *pplug3 pplug4* mutants.



D) Measurement of chloronema length in *pplug3 pplug4* mutants.



E) Measurement of chloronema width in *pplug3 pplug4* mutants.

Lines	Presence	Coverage		Development		Leaflets	
		4 weeks	8 weeks	4 weeks	8 weeks	4 weeks	8 weeks
WT	Yes	Complete	Complete	Adult	Adult	WT	WT
#070	No	—	—	—	—	—	—
#105	No	—	—	—	—	—	—
#106	No	—	—	—	—	—	—
#107	No	—	—	—	—	—	—
#110	No	—	—	—	—	—	—
#211	Yes	Reduced	Complete	Mixed	Adult	WT	WT
#212	Yes	Reduced	Complete	Mixed	Adult	WT	WT
#214	Yes	Reduced	Complete	Adult	Adult	WT	WT

F) Gametophore characteristics of *pplug3 pplug4* mutants.

Figure 15: Phenotypic analysis of the overall plant habit, the size of the gametophyte after four weeks of growth, and over protonema and gametophore growth characteristics. Categories that differ from the WT are colored light grey.
A) General gametophyte parameters. **Genotype**: simplified genotype for both *PpLUG3* and *PpLUG4*, listing the effects of the mutations on the protein sequence. **Habit**: overall, simplified plant habit.
B) Size measurements of the gametophyte area after four weeks of growth, for three gametophytes each. Mean area sizes were calculated, followed by pairwise Student's *t*-tests against the WT size distribution, using Benjamini–Hochberg multiple testing correction and displaying significance values via asterisks or as not significant (ns) (Benjamini & Hochberg, 1995; Student, 1908).
C) Protonema growth parameters. **Presence** and **growth dimensions** of protonema cells: [2D] growth of protonema within agar plates, or [3D] growth of protonema above medium. **Growth pattern**: growth characteristics after four weeks or after eight weeks under long day conditions, and after three weeks under short day conditions: [Free]: free growth of protonema through medium. [Clumped]: protonema cells formed dense clumps of cells with only very short protrusions from periphery. [Outgrowing]: compact clumps of protonema with chains of outward-growing cells. Protonema **cell types** present after four or eight weeks of growth: (Chl)oronema, (cau)lonema, and/or (rhi)zoid cells were observed. Asterisks indicate presence of cells that resembled chloronema or caulonema, but showed visible abnormalities.
D) Measurements of chloronema cell length and **E)** cell width in µm after four weeks of growth. Mean and standard deviation (SD) were calculated and used for pairwise Student's *t*-test analysis against WT to infer significant differences between size distributions. Benjamini–Hochberg correction was used for multiple testing correction.
F) Analysis of gametophore developmental parameters. **Coverage**: gametophyte area covered by gametophores. [Complete]: the whole gametophyte was covered by gametophores. [Reduced]: the gametophyte colonized a larger area via protonema than was covered by gametophores. **Development**: gametophore development. [Adult]: most gametophores were fully developed. [Mixed]: gametophores at the center of the gametophyte were fully developed, while those at the gametophyte periphery were younger and still growing. **Leaflets**: Leaflet size and shape.

4.5.2.2 Reproductive capabilities and sporophyte development are largely independent of PpLUG3 and PpLUG4

Next, phenotyping of the reproductive cycle was performed, by inducing the formation of gametangia, followed by triggering fertilization and documenting the development of the sporophytes. As the development of mature gametophores is a necessary precursor for reproductive development in *P. patens* (Rensing *et al.*, 2020), neither the formation of gametangia nor of sporophytes was observed for any „ball of protonema tissue" mutant lines (Fig. 17A). The „WT similar" lines however exhibited the development of both antheridia and of archegonia, which were not perceptibly different from WT gametangia (Fig. 16A – D and 17A). Similarly, the development and maturation of their sporophytes was comparable to the WT , with the exception of mutant line *pplug3 pplug4* #214 of genotype group **5)**, which exhibited severe developmental delays (Fig. 16E – F and 17B). For WT plants, around 75 % of all gametophores carried mature, brown sporophytes four weeks after fertilization, with an additional 18 % possessing immature green sporophytes. Similar distributions were also observed for the „WT similar" mutant lines #211 and #212, while the developmental distribution for line #214 was significantly altered: There, only 16.7 % of all gametophores carried mature brown sporophytes, with the remaining 83.3 % showing no indication of sporophyte development. This observation indicates a role of PpLUG3 and / or PpLUG4 in the regulation of reproduction and the overall reproductive success, as loss of these proteins can cause reduced sporophyte formation.

However, given that two of the three „WT similar“ mutant lines lack any visible effects on sporophyte development, this reproductive function of PpLUG3 and PpLUG4 is likely an ancillary effect, without forming a critical component of sporophyte initiation as a whole.

No differences between *pplug3 pplug4* mutants and the WT were observed in regards to spore formation and the spore coat (Fig. 16G – H).

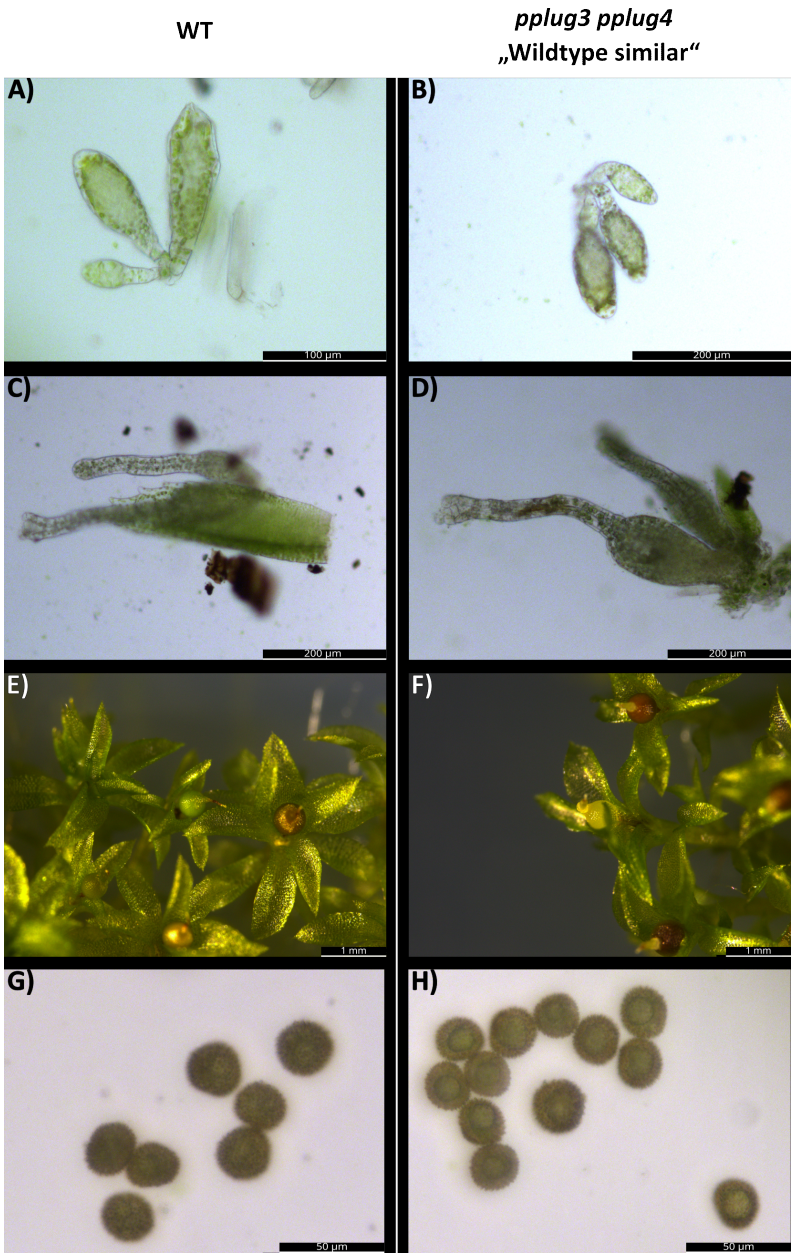
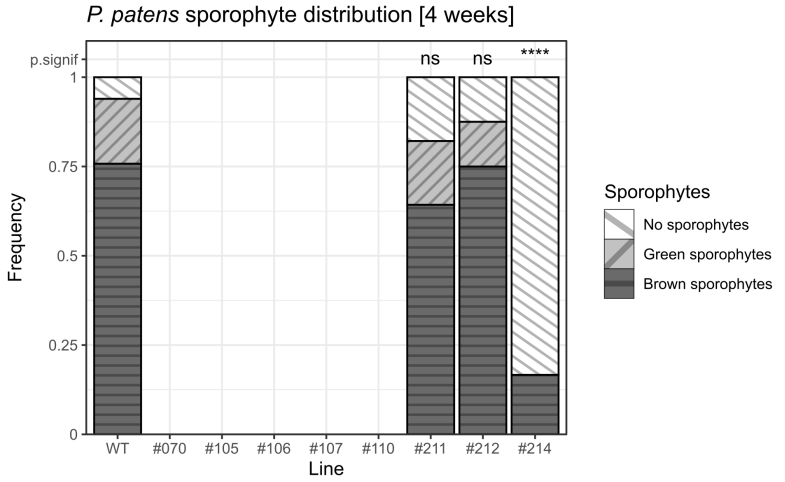


Figure 16: Gametangia and sporophyte development of *pplug3 pplug4 P. patens* mutants. Left column (**A)**; **C)**; **E)**; **G)**): WT plants. Right column (**B)**; **D)**; **F)**; **H)**): *pplug3 pplug4* mutants showing a „WT similar“ phenotype. Depicted are **A)** – **B)** isolated antheridia, **C)** – **D)** isolated archegonia, **E)** – **F)** sporophytes growing on gametophores, and **G)** – **H)** spores. For the *pplug3*ΔLUFS *pplug4* „ball“ mutants, neither gametangia nor sporophytes were present (see Fig. 17).

Lines	Antheridia	Archegonia	Sporophytes	Spores
WT	Present	Present	Present	Present
#070	Absent	Absent	Absent	Absent
#105	Absent	Absent	Absent	Absent
#106	Absent	Absent	Absent	Absent
#107	Absent	Absent	Absent	Absent
#110	Absent	Absent	Absent	Absent
#211	Present	Present	Present	Present, WT looking
#212	Present	Present	Present	Present, WT looking
#214	Present	Present	Present	Present, WT looking

A) Presence or absence of reproductive organs and sporophytes in *pplug3 pplug4* mutant lines.



B) Normalized distribution of sporophyte developmental stages across *pplug3 pplug4* mutants four weeks after fertilization.

Figure 17: Reproductive phenotyping of *pplug3 pplug4 P. patens* mutant lines. A) Classification of [presence] or [absence] of **antheridia** and **archegonia**, **sporophytes**, and **spores** across mutant lines. Categories that differ from the WT are colored light grey. **B)** Normalized sporophyte developmental distribution. All gametophores present were counted and categorized according to sporophyte presence and developmental stage (**no**, **green**, or **brown** sporophytes), and frequencies of each category were calculated. Mutant lines without gametophores were not included. For the sporophyte distribution, pairwise Fisher's exact test was employed, to test for differences between frequency distributions. Benjamini–Hochberg multiple testing correction was used and significance labels were displayed on top, either as asterisks or not significant (Benjamini & Hochberg, 1995; Fisher, 1970).

4.5.3 PpLUG3 and PpLUG4 function in phytohormone-driven developmental pathways of *P. patens*

The „ball of protonema tissue“ mutant lines bear strong resemblances to auxin and cytokinin insensitive mutants described by Ashton *et al.* (1979), to PIN overexpression lines (Viaene *et al.*, 2014), and to mutants of certain auxin-controlled genes like *PpRSLs* and *PpLRLs* (Jang & Dolan, 2011; Tam *et al.*, 2015). Therefore, the *pplug3 pplug4* mutant lines were grown for four weeks on KNOP-ME medium containing high (5 µM) amounts of the synthetic auxin NAA, or of the cytokinin 6-BAP, to illuminate their possible connection to these phytohormone signaling pathways in *P. patens*.

4.5.3.1 The „ball of protonema tissue“ *pplug3*ΔLisH *pplug4* mutant lines are largely insensitive to exogenous auxin

When grown on high concentrations of exogenous auxin, WT plants exhibited a strongly enhanced growth of protonema tissue and a complete loss of gametophore initiation and growth (Fig. 18A, D, and 19A). Most of the present protonema consisted of pigmented caulonema, resulting in an overall notable brownish green coloration, instead of the normal green color of *P. patens*, which was in line with previous reports of the effects of high NAA concentrations (Ashton *et al.*, 1979). In contrast to untreated *P. patens* WT, protonema grew freely in all dimension, even above the medium, and were also induced on top of the initially transplanted gametophore, most likely in the form of rhizoids (Fig. 19B; Lavy *et al.*, 2016). Given the structural similarities between caulonema and rhizoids, it was not possible to differentiate between these two protonema types, especially due to the auxin-induced coloration of caulonema (Ashton *et al.*, 1979). Independently of specific identities however, protonema were generally misshapen, often exhibiting bumps and ridges, as well as bends and sharp turns (Fig. 18D). Additionally, chloronema grown on NAA were significantly shorter by around 10 to 20 %, and also wider by a few µm (Fig. 19C – D).

The „WT similar“ mutant plants behaved largely the same to the WT, also showing an exclusive, 3D growth of protonema, which were often misshapen and consisted of an especially high amount of caulonema or rhizoids, and a brownish green color (Fig. 18B, E, and 19). Further, chloronema cells were significantly shorter than without exogenous auxin, and oftentimes a bit wider, albeit with a low degree of significance (Fig. 19C – D). When comparing the size distributions of auxin-treated WT chloronema against those of the „WT similar“ mutant lines, only the chloronema of mutant line #214 showed significant differences in length and width to the WT chloronema, while those of lines #211 and #212 behaved identical to the WT (Supplementary Table 2).

The „ball of protonema tissue“ phenotype lines however formed a stark contrast to the WT and the „WT similar“ mutant lines, as they appeared to be largely insensitive to exogenous NAA (Fig. 18C, F). They exhibited their previously described „ball“ phenotype,

consisting of densely packed chloronema-like cells with only short protrusion of chains of cells from the ball periphery, and the usual green color of *P. patens* plants (Fig. 19A – B). Caulonema, rhizoids, and gametophores were absent from these lines as well. Nevertheless, there were auxin-induced effects on the growth of chloronema present: For one, chloronema cells appeared often misshapen, as described for WT protonema, and their size also deviated from that of untreated chloronema. Interestingly however, while auxin-grown WT chloronema grew significantly shorter and wider than untreated ones, the chloronema cells of the „ball" phenotypes showed an inverse behaviour, instead growing longer than without auxin, and often a bit narrower, too (Fig. 19C – D). Overall, while the chloronema of WT and „ball" phenotypes were highly divergent under normal growth conditions, they became more similar to each other under the influence of exogenous auxin, with especially the width of most „ball" mutant chloronema showing no significant difference to the width of auxin-treated WT chloronema (Supplementary Table 2).

In summary, the „ball of protonema tissue" *pplug3*ΔLisH *pplug4* mutant lines were largely insensitive to exogenous auxin, indicating that PpLUG3 and PpLUG4 play an important role during the auxin-mediated transcriptional regulatory pathway, and especially for the differentiation of caulonema identity (Jaeger & Moody, 2021; Thelander *et al.*, 2018). As the complete knockout of *PpLUG3* and *PpLUG4* behaved similarly to the WT, some redundancy between all four *P. patens* LUG homologs within this pathway is suggested. However, given the modulation of chloronema shape by exogenous auxin for both principal phenotypes, the LUG homologs are most likely not connected to acidic cell growth via the transcriptionally independent rapid auxin response (Kumar *et al.*, 2025).

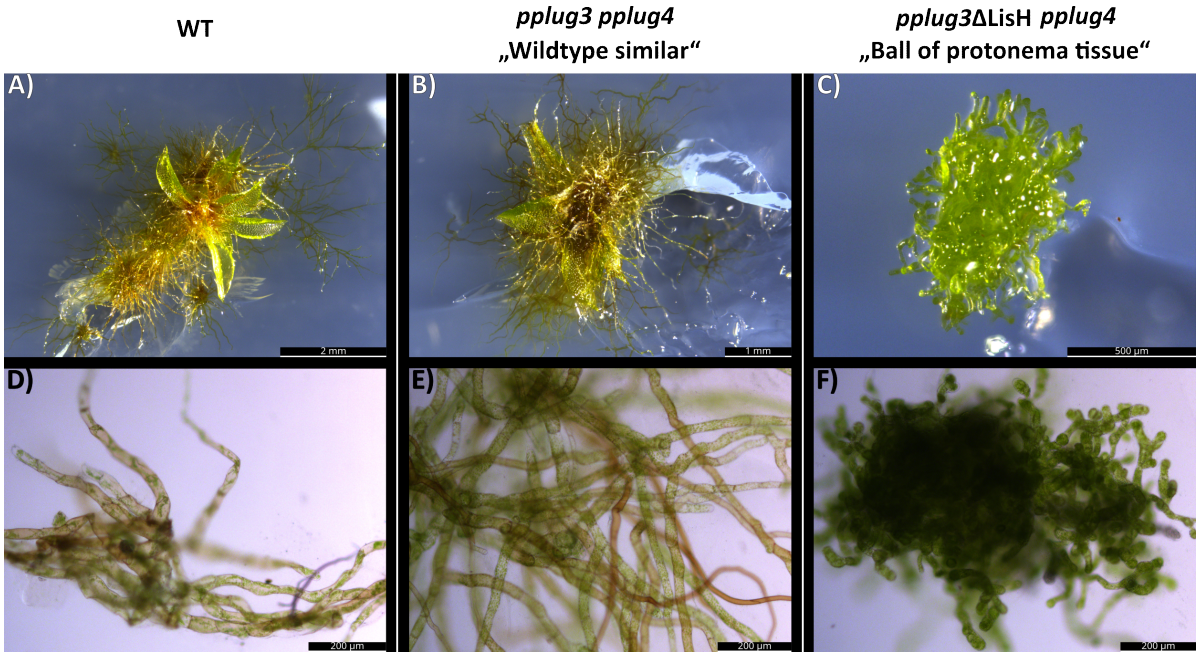


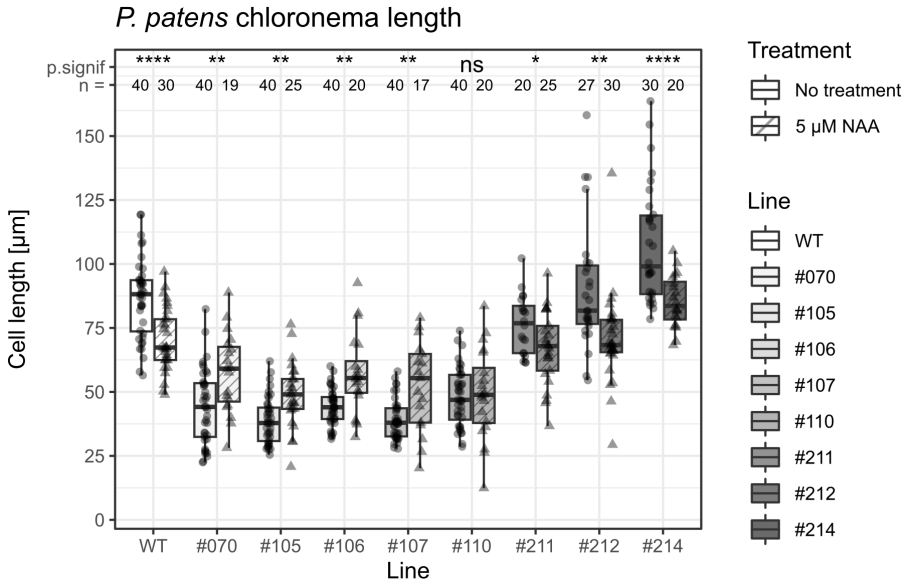
Figure 18: The effect of 5 µM exogenous NAA on *pplug3 pplug4 P. patens* mutants. Left column (**A)**; **D)**): WT plants. Middle column (**B)**; **E)**): *pplug3 pplug4* mutants showing a „WT similar" phenotype. Right column (**C)**; **F)**): *pplug3*ΔLUFS *pplug4* mutants resembling „balls of protonema tissue". Depicted are **A)** – **C)** the whole gametophyte and **D)** – **F)** protonema cells.

Lines	Phenotype		Gametophores
	Observations	Coloration	
WT	Exclusive growth of protonema	Brownish green	Only transplanted present
#070	Similar to untreated „balls of protonema tissue“	Green	Absent
#105	Similar to untreated „balls of protonema tissue“	Green	Absent
#106	Similar to untreated „balls of protonema tissue“	Green	Absent
#107	Similar to untreated „balls of protonema tissue“	Green	Absent
#110	Similar to untreated „balls of protonema tissue“	Green	Absent
#211	Exclusive growth of protonema	Brownish green	Only transplanted present
#212	Exclusive growth of protonema	Brownish green	Only transplanted present
#214	Exclusive growth of protonema	Brownish green	Only transplanted present

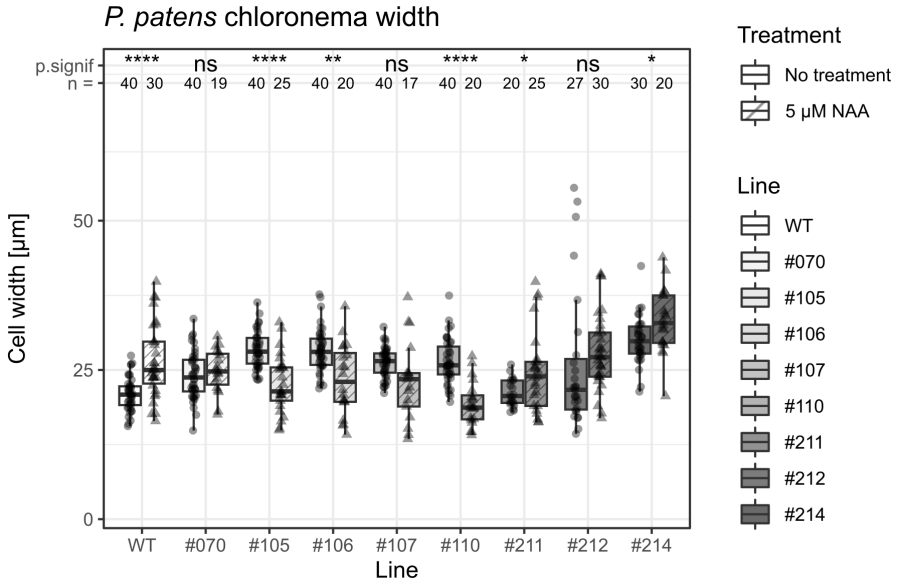
A) General phenotypes of *pplug3 pplug4 P. patens* mutants treated with 5 µM NAA.

Lines	Protonema				
	Growth	Distribution	Types	Form	Notes
WT	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen	Elevated amount of caulonema / rhizoids
#070	3D growth	Clumped	Chl.*	Uneven, misshapen	—
#105	3D growth	Clumped	Chl.*	Uneven, misshapen	—
#106	3D growth	Clumped	Chl.*	Uneven, misshapen	—
#107	3D growth	Clumped	Chl.*	Uneven, misshapen	—
#110	3D growth	Clumped	Chl.*	Uneven, misshapen	—
#211	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen	Elevated amount of caulonema / rhizoids
#212	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen	Elevated amount of caulonema / rhizoids
#214	3D growth	Free	Chl.*; Cau. / Rhi.	Uneven, misshapen	Elevated amount of caulonema / rhizoids

B) Protonema characteristics of *pplug3 pplug4 P. patens* mutants treated with 5 µM exogenous NAA.



C) Measurement and comparison of chloronema length in *pplug3 pplug4* mutants grown with or without 5 µM NAA.



D) Measurement and comparison of chloronema width in *pplug3 pplug4* mutants grown with or without 5 µM NAA.

Figure 19: Phenotypic analysis of *pplug3 pplug4 P. patens* mutants treated with 5 µM NAA. Characteristics differing from WT are colored light grey**.**
A) General **phenotype** and **gametophore** parameters of plants grown on 5 µM exogenous auxin. Mutants either exhibited an exclusive growth of protonema cells and a brownish green colour, or resembled untreated „balls of protonema tissue" (Fig. 15). No growth of new gametophores beyond those transplanted on NAA-containing medium was observed.
B) Protonema characteristics of auxin-grown plants. **Growth**: [3D] growth of protonema above medium. **Distribution**: [Free]: free growth of protonema through medium. [Clumped]: protonema cells formed dense clumps of cells with only very short protrusions from periphery. Protonema **cell types** present after four weeks of growth: (Chl)oronema, (cau)lonema, and/or (rhi)zoid cells were observed. Asterisks indicate presence of cells resembling chloronema or caulonema, but showed visible abnormalities. **Form**: Protonema cells were generally slightly misshapen.
C) Measurements of chloronema cell length and **D)** width for chloronema grown without exogenous auxin (left boxplots, see Fig. 15D and 15E, respectively), and in presence of 5 µM NAA (right boxplots, striped), for each mutant line. Significance values were calculated using Student's *t*-test between the two treatments within each mutant line, and corrected using Benjamini–Hochberg multiple testing correction. Significance levels were displayed on top, either as asterisks or not significant (Benjamini & Hochberg, 1995; Student, 1908).

4.5.3.2 The principal *pplug3 pplug*4 phenotypes differ in the formation of leaflets from cytokinin-induced callus-like clusters

The presence of exogenous cytokinin caused the proliferation of callus-like tissue clusters throughout the WT plants, as described previously (Fig. 20A, D – E; Ashton *et al.*, 1979). These callus-like clusters consisted of clumps of densely packed cells that notably deviated from the structure of protonema cells, as they were small, roundish cells, often exhibiting only few chloroplasts. These clusters were growing from both protonema and gametophore tissue with a high frequency, and often contained outgrowth of gametophore leaflets lacking midribs (Fig. 21). Indeed, these callus-like clusters have been suggested to replace gametophore buds under high cytokinin concentrations, consisting of ectopic stem cells that fail to form stems, but often develop modified leaflets directly from the clusters (Ashton *et al.*, 1979; Cammarata *et al.*, 2023). As such, it was difficult to determine if the plants exhibited multiple gametophores, or merely the outgrowth of leaflets from callus-like clusters. Regarding protonema, both chloronema and caulonema were present, but neither type was noticeably enriched within the gametophyte. However, it was not possible to differentiate between caulonema and rhizoids (Fig. 21A). Protonema were generally freely growing across the gametophyte in a 3D fashion, and were often noticeably malformed (Fig. 20E). Additionally, due to technical difficulties and the prevalence of callus-like clusters, the growth parameters of chloronema cells couldn't be measured.

For the „WT similar" *pplug3 pplug4* mutant lines, both the high proliferation of callus-like clusters and the strong reduction of gametophore formation were observed. In contrast to WT plants however, only the transplanted gametophores were present, and no formation of other possible gametophores was discovered (Fig. 20B, F, and 21). Both chloronema and caulonema were present and growing freely, albeit slightly misshapen; and the outgrowth of leaflets lacking midribs from callus-like clusters was observed as well.

While the „balls of protonema tissue“ lines were affected by exogenous cytokinin, they retained some of their usual characteristics (Fig. 20C, G, and 21): As such, they still exhibited their distinctive „ball“ phenotype and consisted mainly of clumped chloronema cells with only short protrusions from the ball periphery, and a complete lack of both caulonema and rhizoid cells. However, clusters of callus-like cells were interspersed both within the core of the ball, and at its periphery. Additionally, the outgrowth of leaflets was noticeably absent from these callus-like clusters, indicating a critical difference between the „ball of protonema tissue“ phenotypes and the „WT similar“ and WT lines.

In conclusion, while all tested mutant lines were sensitive to the effects of exogenous cytokinin within their usual phenotypes, there were notable differences between the „WT similar“ and „ball“ mutant lines present, namely in the lack of leaflet outgrowth from callus-like tissue clusters. As the formation of leaflets from cytokinin-induced stem cells has been shown to depend on auxin as well (Cammarata *et al.*, 2023), these results corroborate the previous hypothesis of a role of PpLUG3 and PpLUG4 in the auxin signaling pathway of *P. patens*.

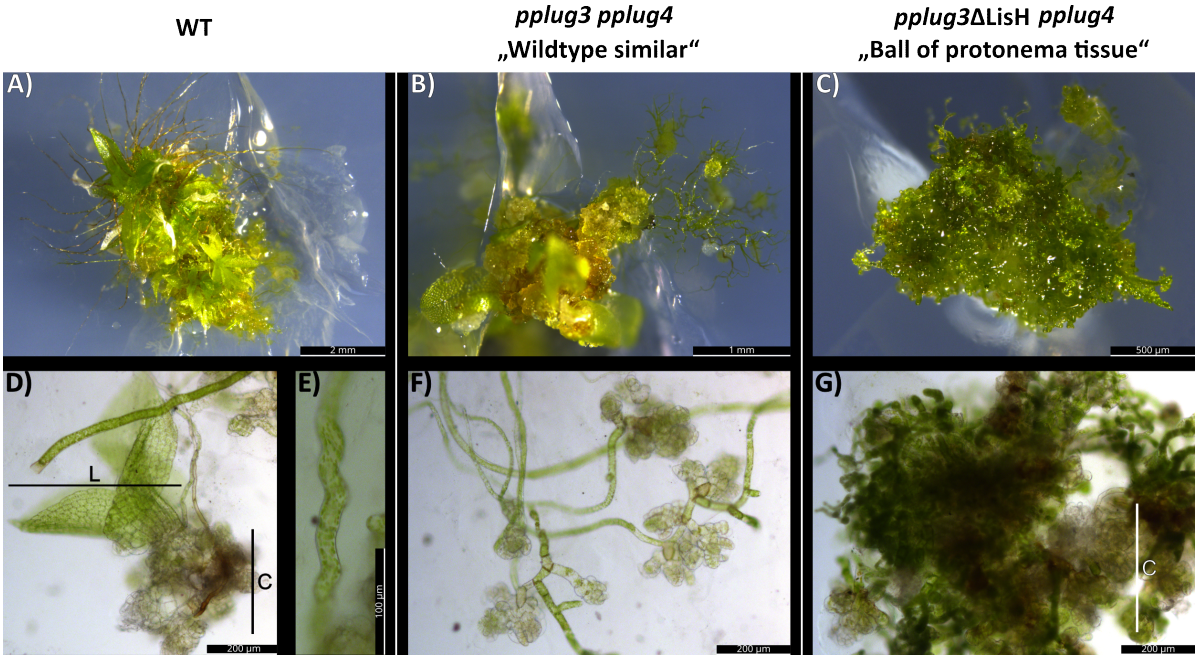


Figure 20: The effect of exogenous 6-BAP on *pplug3 pplug4 P. patens* mutant lines. Left column (**A)**; **D)** – **E)**): WT plants. Middle column (**B)**; **E)**): *pplug3 pplug4* mutants showing a „WT similar“ phenotype. Right column (**C)**; **F)**): *pplug3*ΔLUFS *pplug4* mutants resembling „balls of protonema tissue“. Depicted are **A)** – **C)** the whole gametophyte and **D)** – **G)** protonema cells, clusters of (C)allus-like tissue, and outgrowth of (L)eaflets.

Lines	Phenotype	Protonema			
		Growth	Distribution	Types	Form
WT	Strong proliferation of callus-like tissue across gametophyte	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen
#070	„Balls of protonema tissue“ with callus inclusions	3D growth	Clumped	Chl.*	Uneven, misshapen
#105	„Balls of protonema tissue“ with callus inclusions	3D growth	Clumped	Chl.*	Uneven, misshapen
#106	„Balls of protonema tissue“ with callus inclusions	3D growth	Clumped	Chl.*	Uneven, misshapen
#107	„Balls of protonema tissue“ with callus inclusions	3D growth	Clumped	Chl.*	Uneven, misshapen
#110	„Balls of protonema tissue“ with callus inclusions	3D growth	Clumped	Chl.*	Uneven, misshapen
#211	Strong proliferation of callus-like tissue across gametophyte	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen
#212	Strong proliferation of callus-like tissue across gametophyte	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen
#214	Strong proliferation of callus-like tissue across gametophyte	3D growth	Free	Chl.; Cau. / Rhi.	Uneven, misshapen

A) General phenotype of *pplug3 pplug4 P. patens* mutants treated with 5 µM exogenous 6-BAP .

Lines	Gametophores	Callus-like tissue		
		Present	Tissue structure	Outgrowth
WT	Likely multiple gametophores	Yes	Small, round, densely packed cells	Outgrowth of leaflets lacking midribs
#070	Absent	Yes	Small, round, densely packed cells	No outgrowth
#105	Absent	Yes	Small, round, densely packed cells	No outgrowth
#106	Absent	Yes	Small, round, densely packed cells	No outgrowth
#107	Absent	Yes	Small, round, densely packed cells	No outgrowth
#110	Absent	Yes	Small, round, densely packed cells	No outgrowth
#211	Only transplanted	Yes	Small, round, densely packed cells	Outgrowth of leaflets lacking midribs
#212	Only transplanted	Yes	Small, round, densely packed cells	Outgrowth of leaflets lacking midribs
#214	Only transplanted	Yes	Small, round, densely packed cells	Outgrowth of leaflets lacking midribs

B) Gametophore and callus characteristics of *pplug3 pplug4 P. patens* mutants treated with 5 µM 6-BAP.

Figure 21: Phenotypic analysis of *pplug3 pplug4 P. patens* mutants treated with 5 µM 6-BAP. Characteristics differing from WT are colored light grey**.**
A) General **phenotype** and **protonema** characteristics of mutants grown on 5 µM cytokinin. Two principal **phenotypes** were present: plants with strong growth of callus-like tissue across the whole gametophyte, or „balls of protonema tissue" with inclusions of callus-like clusters. **Growth**: [3D] growth of protonema above medium. **Distribution**: [Free]: free growth of protonema through medium. [Clumped]: protonema cells formed dense clumps of cells with only very short protrusions from periphery. Protonema **cell types** present after four weeks of growth: (Chl)oronema, (cau)lonema, and/or (rhi)zoid cells were observed. Asterisks indicate presence of cells resembling chloronema or caulonema, but showed visible abnormalities. **Form**: Protonema cells were generally slightly misshapen.
B) Characteristics of **gametophores** and **callus-like** clusters. [Multiple gametophores]: Growth of additional gametophores besides the transplanted one. [Only transplanted]: only the transplanted gametophore was present, no further induction of more gametophores observed. **Tissue structure**: Overall phenotype of callus-like cluster and of cells within the clusters. **Outgrowth:** Some mutant lines showed outgrowth of leaflet organs from callus-like clusters.

5 Discussion

5.1 The interaction between the LUG and SEU families coevolved by reciprocal selective pressure since at least 800 million years ago

Peptides and proteins are one of the earliest fundamental aspects of life, and as such function as corner stones of living system in concert with other proteins, DNA, RNA, or other factors like signaling molecules or nutrients. Their interplay forms the foundation of biological complexity and variability, with especially specificity, localization, and timing of interactions being a critical component across all aspects of life (Konieczny *et al.*, 2023). The evolution of these interactions between various factors plays an important role in understanding the ancestral state and the evolutionary trajectory of such interactions, as well as their extant functions and capabilities. Oftentimes, evolutionary changes in one interacting element enable or force cascading changes in another element, which has been described as „reciprocal evolutionary changes in evolutionary interacting loci", or coevolution between elements. Alternatively, high reciprocal selective pressure between elements can force a strong coevolution as well, keeping both elements conserved and their interaction maintained across vast evolutionary distances (Lovell & Robertson, 2010; Thompson, 1994). Especially for protein – protein interactions, it has been assumed that the overwhelming majority of protein sites can't readily accept amino acid substitutions at any given time without prior permissive changes at other loci (Povolotskaya & Kondrashov, 2010), and that especially interaction interfaces of permanent interactions are among the strongest conserved protein regions (Andreani & Guerois, 2014; Lovell & Robertson, 2010; Rodriguez-Rivas *et al.*, 2016).

In the case of the LUG and SEU families of transcriptional regulators, there are multiple pieces of evidence indicating coevolution by strong reciprocal conservation of the core interaction between the two families: The LUG and SEU families are prevalent throughout the streptophyte lineage (Supplementary Fig. 3 and 4), with at least the LUG family having been independently dated to the MRCA of the streptophyte algae, roughly one billion years ago (Bierenbroodspot *et al.*, 2024; Wilhelmsson *et al.*, 2017). It is therefore quite remarkable that the vast majority of analyzed LUG and SEU homologs show a strong similarity between each other in regards to both their canonical domains and in secondary protein folding, even across different sub-clades and between sister lineages (Fig. 5). While there are some variations present, especially the absence of the LisH motif for some algae and bryophyte species, or differences in the LUG C-termini, many of these differences are caused by either incomplete genome annotation of affected species (see section 4.1), or by the positional confidence cutoff used for secondary protein structure annotation (see section 3.2.3). Alternatively, differences might be caused by lineage-specific mutations, as can be seen for example in the bryophyte *P. patens*, where especially PpSEU2 exhibits an

aberrant folding of the LDB-domain. As all species with multiple LUG and SEU homologs exhibit at least one homolog with the preserved canonical LUG or SEU structure, respectively, such differences can potentially be compensated by the function of canonical proteins, thus opening up additional ways of interaction and regulation by allowing sub- and neofunctionalization for these deviating homologs.

Secondly, analysis of the ratio of non-synonymous to synonymous substitutions between algae and land plant *LUG* and *SEU* homologs (Fig. 6) revealed a moderately to strong purifying selection for not only regions encoding functional protein domains ($K_a/K_s < 0.2$), but also for the intermediate region of *LUG* and the IDRs of *SEU*. Even though the high observed K_s rates (Supplementary Table 1) reduce the power of the K_a/K_s analysis by not accurately modeling real substitution rates due to saturation of K_s (Vanneste *et al.*, 2013), the general purifying selection across *LUG* and *SEU* sequences can still be supported: As this behaviour results in an inflation of the K_a/K_s ratio (Anisimova & Yang, 2007; Gharib & Robinson-Rechavi, 2013; Vanneste *et al.*, 2013), the real rate of non-synonymous to synonymous substitutions is likely lower than calculated, and indeed, it has been shown that ratios supporting purifying selection are more robust to K_s saturation (Gharib & Robinson-Rechavi, 2013).

In addition to sequence and structural conservation of the LUG and SEU families, both *in silico* and *in vitro* analyses demonstrated a conservation of their interaction across at least 800 million years of independent evolution as well (Fig. 9 and 10): In all tested major land plant lineages and in Klebsormidiophyceae, LUG and SEU homologs physically interacted with each other, indicating that the formation of the LUG – SEU module originated in the MRCA of these lineages at the latest. Additionally, multiple residues of the LUFS domain could be identified that consistently interacted with SEU homologs residues across the streptophyte lineage (Fig. 7), which together with the strict conservation of the aa residues in this domain and the absence of any preferential interactions between specific phylogenetic clades support the coevolution of this interaction by strong reciprocal conservation. While there were LUG and SEU homologs that didn't show any interaction with each other, lack of dimer formation was probably caused by lineage specific mutations, such as gene or genome duplications, leading to a rapid divergence of one of the duplicated genes. This often causes partial loss of inherited interaction, while the non-diverging homolog usually retains its previous interactions (Andreani & Guerois, 2014; Sun & Kim, 2011). This mechanism is likely especially pronounced in *P. patens*, both in regards of its limited LUG – SEU dimerization capabilities, and in the context of the nearly total absence of PpLUG2 expression (Fig. 8C), as it experienced two ancestral WGDs during the last 70 million years (Bi *et al.*, 2024; Lang *et al.*, 2018).

Taken together, while there were some lineage specific duplications or losses of structural components or interactions for and between the LUG and SEU homologs, the general formation of the LUG – SEU module was highly maintained across at least 800 million years

of streptophyte evolution (Bierenbroodspot *et al.*, 2024). Their evolution is marked by a strong selective pressure to maintain their interaction, which functions as a gene regulatory platform connecting a multitude of different TFs to histone remodels and other transcription regulating processes (Fig. 4; Gonzalez *et al.*, 2007; Sridhar *et al.*, 2004).

However, these observation and their various known roles in *A. thaliana* development raise several questions: For one, how exactly is the module controlled, and what are the underlying dynamics that ensure its correct regulation across these vastly different functions? When did their interaction with the diverse set of TF families and histone modifiers emerge, and are these interactions conserved as well? And lastly, how did the the functions of the LUG – SEU module expand and evolve, especially in connection with their roles in novel gene regulatory networks like in plant stress response (Geng *et al.*, 2017; Shrestha *et al.*, 2014; You *et al.*, 2019), phytohormone signaling (J. E. Lee *et al.*, 2014; Pfluger & Zambryski, 2004), and sexual development in the flower (Azhakanandam *et al.*, 2008; Bao *et al.*, 2010; Conner & Liu, 2000; Franks *et al.*, 2006; Z. Liu *et al.*, 2000; Shi *et al.*, 2024; M. K. Simon *et al.*, 2017; Stahle *et al.*, 2009)?

5.2 The LUG – SEU module represents a conserved protein interaction hub coordinating a variety of different regulatory networks

5.2.1 The breadth and mode of interactions mediated by the module signify its role as a regulatory „date“ hub

One important feature of protein interaction networks is the fact that, due to the dynamics of continuous growth and preferential attachment, older parts of the network are more densely connected than evolutionary younger ones, and a relative small number of proteins are involved in a majority of interactions present (Albert & Barabási, 2002; Barabási, 2013; Batada *et al.*, 2006; Eisenberg & Levanon, 2003; Ramos *et al.*, 2024; Stelzl *et al.*, 2005). Additionally, a negative correlation between the degree of connection and the evolutionary rate of proteins has been described (Fraser *et al.*, 2002). When looking at known protein interactions for AtLUG and AtSEU in *A. thaliana*, even when focusing only on low-throughput studies with a high degree of certainty (Supplementary Table 3), both proteins exhibit a high amount of individual interactions with a variety of different partners, like histone modifiers such as HDA19 (Gonzalez *et al.*, 2007) or SDG4 (Zhai *et al.*, 2020), patterning factors like the YABBY family (Franks *et al.*, 2006; Stahle *et al.*, 2009), or MADS-box TFs (Balanzà *et al.*, 2016; Sridhar *et al.*, 2004). Similarly, even though they are not as well studied as their respective family members, both AtLUH and the AtSLKs exhibit some overlap with, but also distinctive interactions from their homologs, like between AtLUH with histones H2B and H3 (Shrestha *et al.*, 2014) or with PHYTOCHROME-INTERACTING FACTOR 1 (PIF1) (N. Lee *et al.*, 2015).

Interaction networks follow a power law distribution with $f(x) \sim x^{-\alpha}$, where the exponent α is a constant that significantly shapes the distribution (Ng & Huang, 2004). The range of

the exponent has been estimated to span between 1.5 and 2.4 (Ng & Huang, 2004; Stelzl *et al.*, 2005), with some proteome interaction databases showing extreme values of up to 5.3 (Ramos *et al.*, 2024). General approximations of the mean number of interactions each protein in a network possesses have been ranged between two and six, especially when removing the top few interacting proteins, as they tend to inflate the mean significantly (Arabidopsis Interactome Mapping Consortium *et al.*, 2011; Grigoriev, 2003; L.-P. Li *et al.*, 2022; Ng & Huang, 2004). When combining the results of the Y2H assays (Fig. 10 and 12) with the interaction partners known from of low-throughput studies (Supplementary Table 3), AtLUG exhibits 22 known interaction partners, while AtSEU interacts with 29 different proteins. This places both proteins among the top 22 % of interacting proteins in biological interaction networks, when assuming an α value of 1.5, while they place among the top 5 % for higher α values (Fig. 22). However, it is important to note that the curated list of interaction partners is likely not complete, and that not all interactions AtLUG and AtSEU partake in are known, which causes an underestimation of the true amount of interaction partners both proteins possess.

Therefore, the LUG and SEU families represent not only an old, conserved, and coevolving set of transcriptional regulators, but are also highly connected within protein regulatory networks. While a unified definition of hub proteins in interaction networks is still missing, both LUG and SEU can be regarded as such interaction hubs under multiple popular descriptions (Meng *et al.*, 2022).

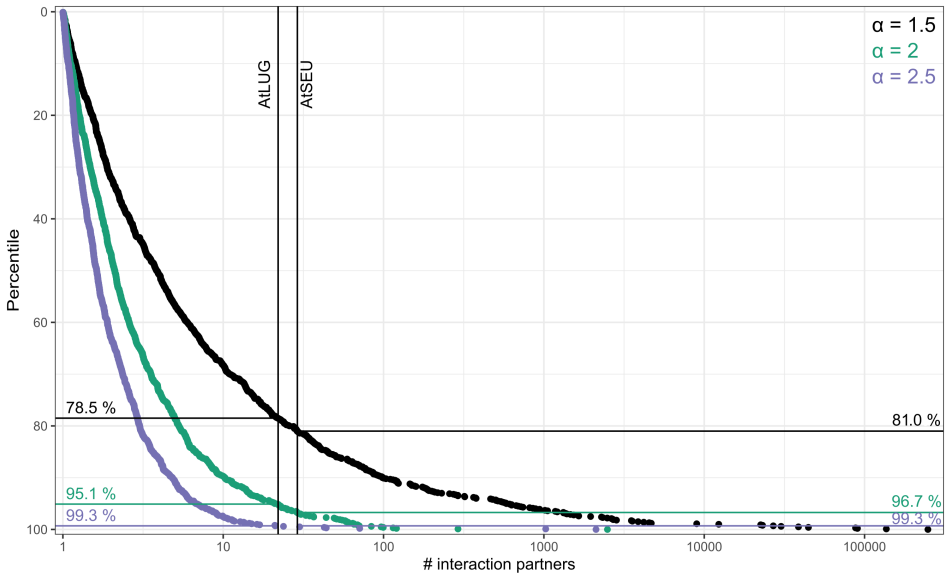


Figure 22: Exemplary power law functions for different α values. X-axis: logarithmic scale representing the values (number of interaction partners) each point (protein) possesses. The y-axis displays the percentile for each point, ranked by decreasing value. 1000 data points were plotted for each function. Black function: power law with α = 1.5; Green function: Power law with α = 2; Purple function: Power law with α = 2.5. Vertical lines depict the position of AtLUG with 22 and of AtSEU with 29 interaction partners, respectively, while horizontal lines depict the percentile of points below these thresholds for the different power law functions. AtLUG percentiles are depicted on the left site, AtSEU on the right site. The functions were created using the poweRlaw package in R (Gillespie, 2015).

Therefore, it is of interest to analyze the general dynamics and control of such protein interaction hubs, and how they apply to the functional LUG – SEU module. Previous studies have identified several behavioral patterns and connected control mechanisms for protein interaction networks as a whole, and for hub proteins specifically. Protein interactions and complexes are generally considered as dynamic states depending on a variety of spatial, temporal, and contextual variations, and thus avenues of control (Meng *et al.*, 2022; Przytycka *et al.*, 2010). They can be controlled by a combination of three distinct mechanisms: By temporal and spatial control of protein encounters, by local concentration regulated via gene expression, degradation, and similar mechanisms, and by the physicochemical environment (Nooren, 2003). Hub proteins in particular possess discrete behaviours and regulatory traits, due to their importance in protein networks: Depending on their characteristics, they are often divided into „party" hubs that possess multiple interaction interfaces and bind most of their interaction partners simultaneously, and into „date" hubs that bind different protein partners at distinct times and locations, often only exhibiting one or two interaction interfaces (Han *et al.*, 2004; P. M. Kim *et al.*, 2006). The interaction partners of such „date" hubs are usually mutually exclusive, and the control of specific interactions relies on precise spatiotemporal expression of specific components (Goel & Wilkins, 2012; P. M. Kim *et al.*, 2006). Two behaviours are prevalent for such hubs, either a dynamically expressed protein hub that relies on competitive binding of their interaction partners to enable distinct interactions, or a static hub whose binding partners are controlled via dynamic expression, but not through mutual competitiveness (Goel & Wilkins, 2012).

When applying these characteristics to the LUG – SEU module, it becomes apparent that these proteins most likely form a „date" hub with static expression, relying on dynamic control of their interaction partners for targeted gene regulation: For one, the examined LUG and SEU homologs showed a high degree of interaction with both types of MADS-box TFs in the Y2H assay, with no apparent preferences for either type (Fig. 12). Their homologs are broadly expressed throughout land plants, with only some selective tissues or developmental stages showing reduced expression, such as *A. thaliana* pollen and embryo, or *M. polymorpha* sperm cells (Fig. 8). Contrastingly, the expression of the target specifying MADS-box genes exhibits a noticeably higher variation in their expression across the various tissues and developmental stages, thus allowing for the formation of different regulatory complexes in diverse contexts (Supplementary Fig. 1). Additionally, both LUG and SEU homologs in *A. thaliana* have been mostly implicated to interact with only one or two proteins at a time, with only a few pathways like the interaction with YABBYs or with the mediator complex exhibiting higher order complexes (Gonzalez *et al.*, 2007; Stahle *et al.*, 2009).

5.2.2 Both LUG and SEU exhibit specific behaviours and structural features supporting their diverse regulatory functions

One aspect highly likely enabling the function of LUG and SEU homologs to act as regulatory hubs are their IDRs. In recent years, research into the characteristics of IDRs in general has revealed their rich functional possibilities, as they allow and govern an immense breadth of roles, from substrate or partner recognition, to abiotic sensoring, and transcriptional regulation (Holehouse & Kragelund, 2024). They often exhibit the ability to dynamically interact with other proteins or DNA in multiple different, mutually exclusive confirmations, called fuzzy binding; and together with their prevalence for post-translational modifications, high association rates, and possibility for small linear motifs increasing affinity for specific protein domains, they represent a critical backbone for regulatory networks, that can be relatively easily fine-tuned and shaped by evolution (Holehouse & Kragelund, 2024; Már *et al.*, 2023; van der Lee *et al.*, 2014). One especially notable ability of IDRs is the formation of phase-separated, liquid-like condensates that can alter gene expression patterns upon forming or dissolution. These condensates are often triggered by abiotic factors like temperature, osmolarity, or mechanical stress, but also by protein concentration, and post-translational modifications (Holehouse & Kragelund, 2024; Már *et al.*, 2023; Xiao & Xia, 2025). For example, formation of such liquid-like condensates has been demonstrated for SEU homologs across land plants, resulting in altered gene expression upon hyperosmotic stress, while SLK proteins are interestingly incapable of condensation (B. Wang *et al.*, 2022).

Of course, there are more structural features and behaviours that aid in the establishment of LUG and SEU as a regulatory hub, with especially the WD40 domain being of note, as it facilitates the formation of larger protein complexes throughout eukaryotes (Xu & Min, 2011). Another example can be seen by the strong degree of homo- and heterodimerization observed within both the LUG and SEU families for multiple species (Fig. 10): It has been previously described that the probability of a protein to self-interact increases linearly with its connectivity in protein interaction networks, and that proteins exhibiting homodimer formation interact with twice as many partners on average than those without self-interaction. This behaviour has been attributed to the „stickiness“ of proteins, which expresses their disposition for forming protein dimers and multimers in general, with a higher „stickiness“ correlating to an increase in homodimer formation (Ispolatov, 2005). Additionally, it has been reported that homologs of homodimer-forming proteins often also display the formation of heterodimers with their family members, due to the inheritance and subsequent modification of homodimer capability, allowing for the binding between different homologs instead (A. Marchant *et al.*, 2019). For the LUG family, homo- and heterodimerization is most likely facilitated by the LisH motif, given its function as general dimerization motif across eukaryotes (Gerlitz *et al.*, 2005; M. H. Kim *et al.*, 2004), and / or by the LUFS domain, which has been suggested to even allow homotetramer formation (H. Wang *et al.*, 2019). While the exact mechanism is still unknown for the SEU

family, it is likely connected to their IDRs, which have been shown to generally modulate both protein interaction as well as dimerization (Már *et al.*, 2023). Therefore, these observed homo- and heterodimers within the LUG and SEU families not only underscore their function as regulatory hubs by indicating their propensity as highly interacting proteins, but could potentially also function in regulatory pathways as well, for example by requiring dimerization ahead of interaction partner binding. However, the possible roles of LUG and SEU homo- and intrafamily heterodimers in their gene regulatory networks remain to be studied.

In conclusion, the connection of LUG and SEU homologs both individually and as a module to a variety of interaction partners in a plethora of regulatory pathways, ranging from reproduction to phytohormone signaling and stress response, are the results of their role as „date“ hubs in gene regulatory networks, and likely modulated by their functional domains, their „stickiness“ for protein interactions, and by operative motifs in and behaviours of their IDRs (Fig. 4; Azhakanandam *et al.*, 2008; Conner & Liu, 2000; J. E. Lee *et al.*, 2014; Shrestha *et al.*, 2014; Sitaraman *et al.*, 2008; B. Wang *et al.*, 2022).

5.3 The recruitment of the functional LUG – SEU module to floral developmental regulation by the MADS-box TFs

The origin of the evolutionary trajectory of the LUG – SEU module as a regulatory platform can be traced back at least 800 million years ago, to the MRCA of the Klebsormidiophyceae and the ZCC-grade streptophyte lineages (Bierenbroodspot *et al.*, 2024): Both the formation of the LUG – SEU module, as well as the interaction of LUG homologs with the MADS-box MIKC family could be dated to this ancestor, indicating a deeply conserved interaction between these proteins. Interestingly, no heterodimer formation could be observed between the *K. nitens* KnSEU and KnMADS homologs, while extensive interactions between their homologs in land plants were present (Fig. 12), indicating either a lineage-specific loss of this interaction for *K. nitens*, or a gain during the evolution of the ZCC-grade and / or the embryophyte lineages. Similarly, some lineage-specific losses between protein binding of the LUG homologs with MIKC^{C} could be observed, like in *C. richardii* and *M. polymorpha*, while interactions between LUG and the MIKC* group were maintained across land plants.

MADS-box genes have been implicated in reproductive processes across both land plants and the ZCC-grade, for example functioning in reproductive cell differentiation in multiple charophyceae species (Tanabe *et al.*, 2005). While little is known about their exact regulatory mechanism in algae or their functions beyond the ZCC-grade, the conserved interaction between LUG and MADS-box homologs suggest that at least some of the MADS-box target genes across streptophyte algae evolution were regulated in concert with LUG proteins. Additionally, multiple families of proteins that are known to act jointly with LUG and / or SEU homologs have been shown to be present across eukaryotes in general, such

as the class I histone deacetylase and the histone acetyltransferase families, to which the *A. thaliana* homologs AtHDA19 and AtHAC1, respectively, belong to (Alinsug *et al.*, 2009; Pandey, 2002). Additionally, multiple components of the mediator complex exist throughout eukaryotes as well (Nagulapalli *et al.*, 2016), while members of the SDG family could be traced to the MRCA of the streptophytes at the latest (Supplementary Fig. 8). Taken together, there is a high probability that the functional gene regulatory complex between target-specifying MADS-box TFs, LUG and possibly SEU homologs, and histone modifiers or other regulatory proteins was already formed during the evolution of the streptophyte lineage, governing aspects of reproductive regulation (Fig. 23). This in turn could have laid the foundation for the subsequent evolution of function of the LUG – SEU module in the land plant lineage.

During land plant evolution, gene regulatory networks were rewired and expanded by gene neo- or subfunctionalization, the occurrence of novel genes, and by co-option of already present regulatory networks, to support the development of new pathways, tissues and organs, such as the auxin signaling pathway, desiccation tolerance mechanisms, or stomata and vascular tissues (Artur & Kajala, 2021; Bowles *et al.*, 2022; Mutte *et al.*, 2018). For example, in the lineage leading to land plants, a split of the MIKC family into the type I and type II MADS-box groups occurred, with the type I family subsequently diverging into the MIKC* and M-type genes (Kaufmann *et al.*, 2005; Kwantes *et al.*, 2012; Qiu *et al.*, 2023; Qiu & Köhler, 2026). Similarly, the emergence of land plant specific clades for both the LUG and SEU families can be detected (Fig. 5). Interactions between SEU homologs with both $MIKC^C$ and MIKC* type proteins could be observed across all tested land plant lineages, as well as interaction with M-type proteins in *P. patens*, without a noticeable bias for either group (Fig. 12). Together with the lineage-specific losses of LUG – $MIKC^C$ interaction mentioned above, it is quite likely that in land plants, the binding of MADS-box TFs to SEU homologs is the primary mechanism of targeting the LUG – SEU module to specific genes, while binding to LUG functions as a secondary option in specific situations (Fig. 23). During the divergence of the MADS-box lineage, $MIKC^C$ proteins became primarily associated with reproductive regulation in sporophytes, while the type I proteins primarily function in regulation of the gametophyte, with MIKC* ones controlling the male gametophyte, and M-types regulating the female gametophyte and the embryo, respectively (Gramzow & Theissen, 2010; Qiu & Köhler, 2022; Qiu & Köhler, 2026; M.-X. Zhang *et al.*, 2020; L. Zhang *et al.*, 2024). Therefore, the recruitment of the LUG – SEU module to multiple aspects of angiosperm reproduction, such as general flower development (Conner & Liu, 2000; Franks *et al.*, 2002; Sridhar *et al.*, 2006) or regulation of ovule initiation and development (Azhakanandam *et al.*, 2008; Bao *et al.*, 2010; Z. Liu *et al.*, 2000; M. K. Simon *et al.*, 2017) was likely mediated by the continued interaction with the MADS-box family, whose diversification and expansion was a critical cornerstone for the evolution of these reproductive structures (Kaufmann *et al.*, 2005; Qiu & Köhler, 2026). Additionally, this also allowed subsequent secondary co-option of LUG and / or SEU functions, for

example in anther development (Shi *et al.*, 2024), and likely aided in the connection between stress responses and the regulation of sexual development (Geng *et al.*, 2017; Qiu & Köhler, 2026; Shrestha *et al.*, 2014).

As the LUG – SEU complex functions as a static „date“ hub, fine-tuning of specific regulatory complexes is likely mediated primarily by the specific expression and binding dynamics of individual MADS-box genes and other protein partners. Multiple MADS-box genes show strong differential expression in various tissues and developmental stages (Supplementary Fig. 1), and additionally, differences in predicted interaction sites between various SEU homologs and selected MIKC^{C} and MIKC* proteins could be observed (Supplementary Fig. 9). Interacting residues in dimers between SEU homologs with both MIKC^{C} members, and the MIKC protein of *K. nitens*, are predicted to be primarily located in the K-domain, while they were distributed across both the I- and K-domain for MIKC* containing dimers. Further, interactions between the CrSEUs of *C. richardii* with the MIKC* protein CRM13, and binding for both MADS-box types in *P. patens* were predicted to involve the M-domain as well. Interactions between MADS-box TFs with both MADS-box and non-MADS-box proteins are thought to be mediated primarily by the K-domain, and to a lesser extend the I-domain (Melzer & Theißen, 2009; Z. Zhang *et al.*, 2024). Therefore, such lineage- and group-specific differences could signify a certain level of independent evolution and fine-tuning between the different groups of MADS-box proteins in regards to their interaction with LUG and SEU homologs, allowing for more precise modulation of binding specificity and target gene regulation. The functions of the more weakly conserved C-domains are less clear, but they have been suggested to play a role in some interactions with other protein classes as well, and indeed, in *A. thaliana*, AtSEP3 and AtAP1 both rely to a certain extend on the presence of their respective C-domains to bind to AtSEU (Melzer & Theißen, 2009; Sridhar *et al.*, 2006; Z. Zhang *et al.*, 2024)

Another possible avenue for controlling the regulatory effects of the LUG – SEU module is the activity and affinity of their interacting histone modulators and other regulatory proteins. Although there is not much known about the exact functional dynamics of these proteins in plants outside angiosperms, there have been reports that the regulatory effects of certain histone modifications depend on the specific plant lineage: For example, H3K4me3 trimethylation, which can be conveyed by the activity of SDG4 homologs, was shown to cause transcriptional repression in *M. polymorpha*, but is associated with activation in *P. patens* (Hisanaga *et al.*, 2023; Zhai *et al.*, 2020). Similarly, single methylation of histone 3 (H3K9me1) in gymnosperms correlates with euchromatin, while it represents a heterochromatin mark in angiosperms (Fuchs *et al.*, 2008). Therefore, the readout of the various histone modifications caused by the LUG – SEU module is another malleable pathway that influences its precise regulatory effects on various target genes.

In summary, the interaction between components of the LUG – SEU module with the MADS-box TFs was maintained for at least 800 million years of evolution in the green lin-

eage (Bierenbroodspot *et al.*, 2024), and likely played an important role in reproductive control in the lineages leading to the ZCC-grade and the embryophytes. After the transition to land, the continued function of MADS-box TFs in all aspects of reproduction anchored the LUG – SEU module to a position as a central regulator of land plant sexual reproduction. Further, this position also aided in the recruitment of the module or of singular components to a variety of different processes, thereby establishing a connection between the regulation of reproduction and additional pathways, such as auxin signaling or stress response.

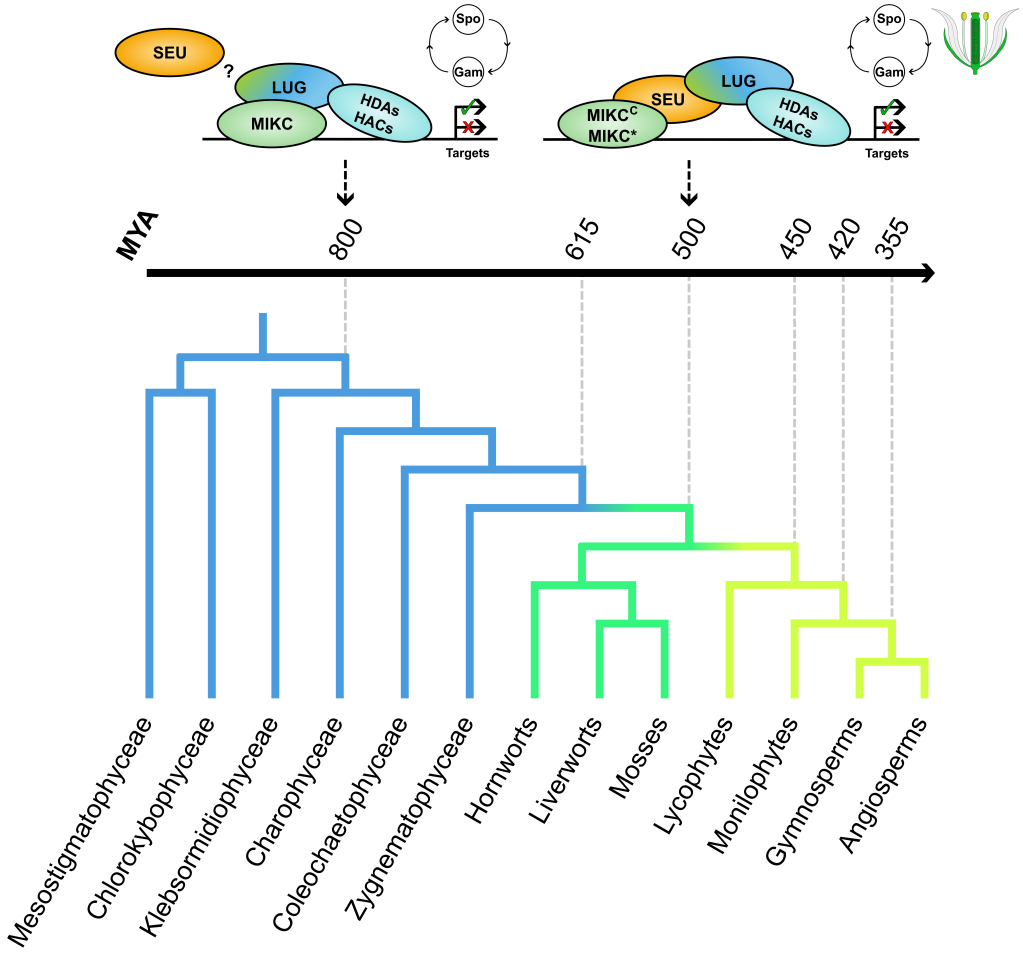


Figure 23: Proposed evolution of the LUG – SEU – MADS-box TF interaction in the streptophyte lineage. Interaction between MIKC type MADS-box TFs, LUG homologs and histone modifiers like HDAs and HACs is suspected to have originated in the MRCA of Klebsormidiophyceae and the ZCC-grade at the latest, governing reproductive processes, and with unclear involvement of SEU homologs. In the lineage leading to land plants, both MIKC^C and MIKC* proteins primarily interact with SEU homologs in the context of the LUG – SEU module. MADS-box TFs are still suspected to govern reproductive regulation, for example flower development in angiosperms. Major lineages for streptophyte algae are colored blue, bryophyte lineages green, and tracheophyte lineages lime (Bierenbroodspot *et al.*, 2024; G.-Q. Liu *et al.*, 2022). Estimated crown ages for selected branches based on Harris *et al.* (2022).

5.4 LUG homologs function in phytohormone signaling and organ identity specification since the origin of land plants

Analyzing the roles of LUG homologs in the developmental regulation of mosses like *P. patens* can not only yield information about the evolution of their functions in the bryophyte lineage, but together with known regulatory pathways from angiosperms, can also aid in reconstructing the ancestral functions of the LUG family during the transition to land 500 million years ago (Harris *et al.*, 2022). Here, it can be shown that LUG homologs function in the regulation of the auxin signaling pathway and of downstream organ specification since the emergence of this pathway, thus functioning in both sexual regulation, but also phytohormone-mediated signaling during plant development.

5.4.1 The *P. patens pplug* phenotypes exhibit a bistable switch between the two principal phenotypes

The genotype – phenotype combinations of the *P. patens pplug3 pplug4* mutants can be divided into four different categories (section 4.5.1): There were two double *pplug3 pplug4* knockouts exhibiting a „WT similar“ phenotype (groups **4)** and **5)**), four individual lines possessing a deletion in the PpLUG3 LUFS domain and a full *pplug4* knockout that displayed a „ball of protonema tissue“ phenotype (group **2)**), as well as two intermediate mutants. Of those, line #070 of group **1)** lacked fragments of both the PpLUG3 and the PpLUG4 LUFS regions, resulting in a „ball“ phenotype, while line #211 (group **3)**) consisted of a similar genotype to **2)**, but exhibited the „WT similar“ phenotype instead. Together, these results suggest that partial deletions in the LUFS interaction domain of especially PpLUG3, and to a lesser extend of PpLUG4, are the main driver of the „ball“ phenotype, while complete knockouts of both *PpLUG3* and *PpLUG4* result in a „WT similar“ phenotype. However, line #211, which possessed a „WT similar“ phenotype, but unexpectedly, a *pplug3*ΔLUFS *pplug4* genotype, demonstrates that the exact connection between *PpLUG* genotype and phenotype is more complex than initially surmised.

To precisely explain the precise mechanistic foundation of the observed phenotypes, four assumptions about LUG homologs in *P. patens* were made: For one, it was postulated that both PpLUG3 and PpLUG4 possess gene regulatory functions in *P. patens*, as the LUG family is not known to act outside of gene regulation (Gonzalez *et al.*, 2007; Sridhar *et al.*, 2004; Vega-León *et al.*, 2025). Secondly, the general functions and therefore also the simultaneous loss of *PpLUG3* and *PpLUG4* can be compensated by activity of PpLUG1. In contrast to PpLUG2, it is not only the most strongly expressed LUG homolog in *P. patens* (Fig. 8), but also showed protein interactions with PpSEU2 and with some MADS-box TFs in the Y2H assays, which were absent for PpLUG2 (Fig. 10 and 12). Additionally, at least partial compensation of the loss of LUG proteins by other LUG homologs has been described previously, for example for *A. thaliana* (Sitaraman *et al.*, 2008). Thirdly, it was assumed that mutations of the LUFS domain primarily result in disruption of the LUG –

SEU complex, as the LUFS domain is not only necessary for interaction with SEU homologs in plants (Shrestha *et al.*, 2014; Sridhar *et al.*, 2004), but also with LDB domain containing proteins across eukaryotes (Bao *et al.*, 2010; van Meyel *et al.*, 2003). Indeed, the deletion observed for the L3_2 mutant for example causes a loss of several of the short N-terminal α-helices of the LUFS domain, which partially overlaps with the region of consistently interacting residues between LUG and SEU homologs (Fig. 7 and 13F). As other proteins bind LUG mostly via the intermediate region or the WD40 domains however (Gonzalez *et al.*, 2007; Vega-León *et al.*, 2025; Xu & Min, 2011), their interactions are unaffected. As discussed in section 5.3, multiple components known to interact with LUG and / or SEU homologs were already present and, in the case of the MADS-box TFs, shown to interact in the MRCA of land plants and the Klebsormidiophyceae lineage. Therefore, it is reasonable to, fourthly, assume that the general mechanism of interaction and target gene regulation by the LUG – SEU complex in the bryophyte lineage is comparable to the pathway known in angiosperms.

Combining all four assumption, the following model for target gene regulation by PpLUG3 and PpLUG4 in *P. patens* is proposed, using the canonical regulation in angiosperms as a basis (Fig. 24): Under normal conditions, PpLUG3 and / or PpLUG4 can interact with PpSEU proteins to bring TFs specifying genetic targets together with histone deacetylases or other histone modifiers (Fig. 24A; Gonzalez *et al.*, 2007; Sridhar *et al.*, 2004; Sridhar *et al.*, 2006). Although no interaction between PpLUG3 and the two tested PpSEU homologs has been observed in the Y2H assay (Fig. 10), the general conservation of the LUG – SEU interaction and the presence of strong phenotypes for *P. patens pplug3* mutants suggest interactions between PpLUG3 with at least some of the PpSEU homologs present. Additionally, Y2H assays are prone to false negative results that can't easily be detected (Brückner *et al.*, 2009), further underscoring a possibility for PpLUG3 – PpSEU interactions.

The loss of both *PpLUG3* and *PpLUG4* causes the remaining PpLUG homologs, primarily PpLUG1, to compensate the loss by acting in the LUG – SEU module in their stead (Fig. 24B). However, disruption of only the LUFS domain of PpLUG3 and / or PpLUG4 causes competitive binding between the mutated and the remaining PpLUG homologs to their shared interaction partners (Fig. 24C). In combination with the inability of mutated PpLUG3/4 to bind to PpSEU, a reduction of the overall efficiency of the PpLUG mediated gene regulation occurs, causing an increased expression of normally repressed genes, and vice versa. Subsequently, downstream regulatory pathways are compromised, causing a cascade effect that results in a switch from the „WT similar" configuration of *P. patens* to the severely disrupted „ball" phenotype. Such bistable switches between two highly stable, but fundamentally different states are quite common throughout biological systems, like for enzymatic proteins or positive feedback loops. Perturbations of such systems, for example by changes in transcriptional regulation, can cause the stability and balance of stable states to change, or can even outright eliminate certain states, thus pushing the system into a different state of stability (D. Huang *et al.*, 2012; Kadiyala *et al.*, 2025). In

the case of *P. patens pplug3 pplug4* mutants, there seem to be two principal stable states, namely the WT state, and the „ball of protonema tissue“ state. Depending on the competitive strengths between intact PpLUG1 and the mutated PpLUG3 and PpLUG4, phenotypes fall into one of these two stable states (Fig. 24D).

This model can also explain the paradoxical genotype – phenotype combination of line #211, and some phenotypic observations of line #214: For the former, even though only the LUFS domain of PpLUG3 was disrupted, with the remaining protein staying intact, a „WT similar“ phenotype was present, indicating that the intact PpLUG homologs might have been competitively stronger than the mutated PpLUG3, thus being able to compensate the compromised PpLUG mediated gene regulation at a higher level than in other mutant lines, preventing a switch to the „ball“ mutant phenotype by adding stability to the WT state. Similarly, while both *PpLUG3* and *PpLUG4* were lost in line #214, resulting in a „WT similar“ phenotype, it behaved differently than the other two „WT similar“ lines: For example, line #214 exhibited a reduced reproductive success, with a strong decrease in sporophyte formation compared to the other mutant lines (Fig. 17B). Additionally, their chloronema length and width were significantly different from the WT, under both normal growth conditions, and with exogenous auxin (Fig. 15D – E, Supplementary Table 1). These observations indicate a reduced ability of PpLUG1 to compensate the loss of *PpLUG3* and *PpLUG4*, resulting in some regulatory downstream effects, but not strong enough to overcome the region of high instability between the „WT similar“ and the „ball of protonema tissue“ phenotype.

However, the exact reasons one or the other PpLUG homolog might be more successful in recruiting their binding partners or to compensate for gene loss remains unknown, but could yield an interesting avenue for further research into the mechanistic details of the LUG – SEU module gene regulation in *P. patens.*

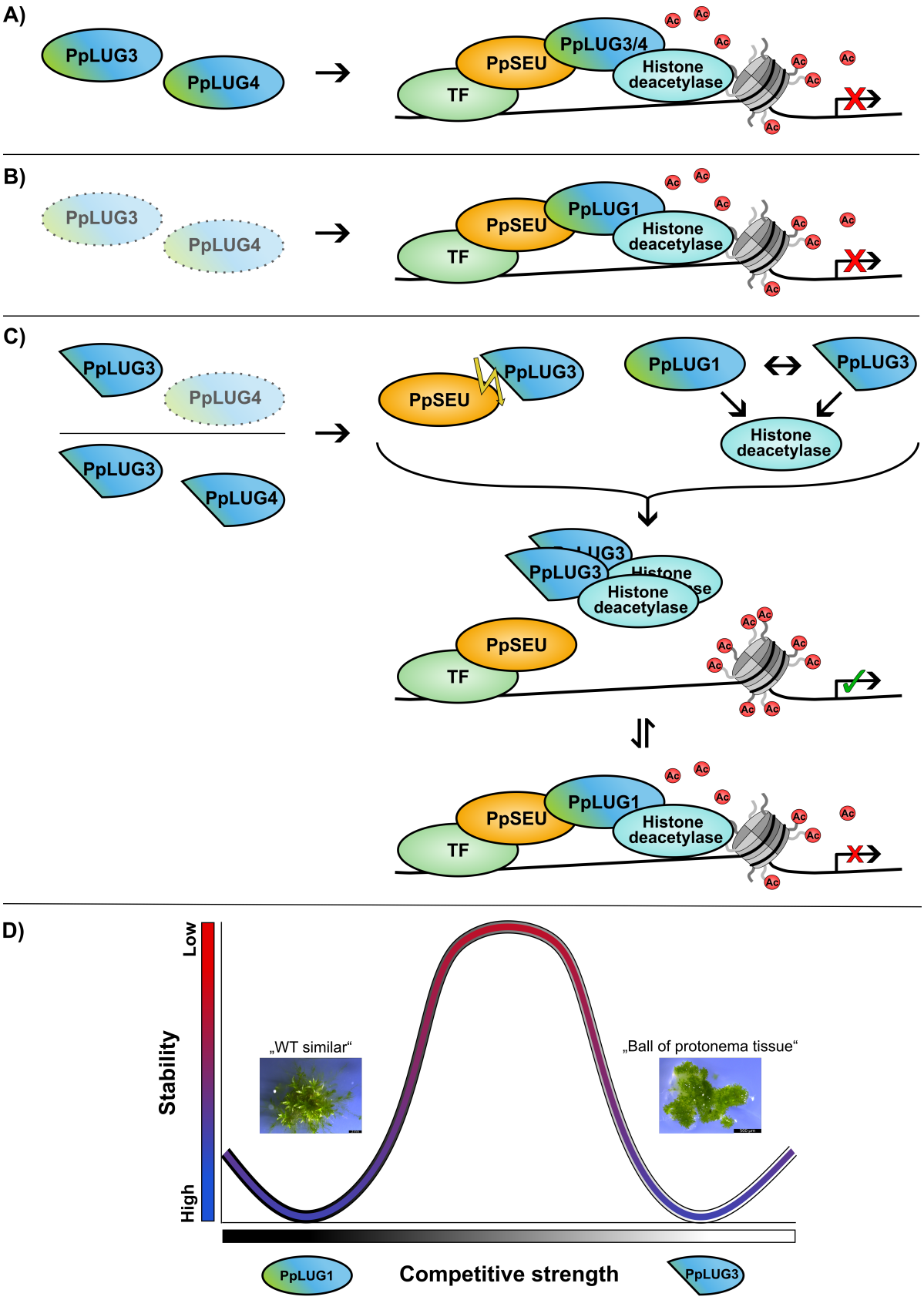

A)
PpLUG3
PpLUG4
PpLUG3/4
PpSEU
TF
Histone deacetylase
Ac
B)
PpLUG3
PpLUG4
PpLUG1
PpSEU
TF
Histone deacetylase
Ac
C)
PpLUG3
PpLUG4
PpLUG3
PpLUG4
PpSEU
PpLUG3
PpLUG1
PpLUG3
Histone deacetylase
PpLUG3
Histone deacetylase
PpSEU
TF
Ac
PpLUG1
PpSEU
TF
Histone deacetylase
Ac
D)
Low
High
Stability
„WT similar“
„Ball of protonema tissue“
PpLUG1
Competitive strength
PpLUG3

Figure 24: Proposed mechanistical model for the *pplug3 pplug4 P. patens* mutant phenotypes. A) Function of the LUG – SEU complex in WT plants, causing deacetylation and repression of target gene expression. **B)** Compensation of the simultaneous loss of PpLUG3 and PpLUG4 in knockout lines by PpLUG1. **C)** Disruption of the LUFS domain for PpLUG3 and loss of PpLUG4 or its LUFS domain prevents formation of the LUG – SEU complex. In combination with competitive binding of LUG interaction partners like histone deacetylases to PpLUG3/PpLUG4 and to intact PpLUG1, the effectiveness of the target gene regulation of the LUG – SEU module is weakened, allowing for expression of otherwise repressed genes, disrupting the stable state. **D)** Model of the bistability of *pplug3 pplug4* mutant lines. Both the „WT similar" and the „ball of protonema tissue" phenotype are stable states divided by a region with low stability. Depending on the competitive strength of intact PpLUG1 and PpLUG3ΔLUFS, the mutant phenotypes fall into one or the other state.

5.4.2 PpLUG homologs are critical for establishing caulonema identity in concert with the auxin signaling pathway

The *P. patens* „ball of protonema tissue" mutant lines primarily consisted of tissue with high similarity to chloronema, but aberrant cell sizes and shapes (Fig. 15C – E). Given that chloronema are the first emergent tissue type in the *P. patens* life cycle, and that the observed cells shared several critical similarities with chloronema, like the signature transverse cell plates and the high amount of chloroplasts (Jaeger & Moody, 2021; Moody, 2022; Rensing *et al.*, 2020), it is reasonable to assume that the primary tissue type of „ball" mutants was indeed comprised of chloronema cells. Under prolonged growth, some „ball" phenotype lines exhibited the formation of protonema cells of questionable caulonema identity, and considering that the transition from chloronema to caulonema in general is only established during the development of the gametophyte (Moody, 2022; Rensing *et al.*, 2020), it is highly likely that the observed phenotype is caused by a severe disruption of either caulonema identity specification, or protonema differentiation. Indeed, three sets of mutants that resemble many aspects of the „ball" mutant phenotypes have been described so far, and all three are connected to the establishment of the chloronema – caulonema transition, which is governed by the auxin signaling pathway:

Double knockouts of the bHLH TF genes *PpRSL1* and *PpRSL2* result in small gametophytes with considerably shorter chloronema cells and an inability to form caulonema, and subsequently gametophores (Jang & Dolan, 2011). The auxin-promoted activity of both *PpRSL1* and *PpRSL2* is required for establishing polar caulonema identity along the protonema filament, and as such, overexpression of *PpRSL1* and *PpRSL2* leads to an increased formation of caulonema cells, resembling WT plants grown on exogenous auxin (Jang & Dolan, 2011). In a similar vein, knockouts of the *PpLRL1* and *PpLRL2* genes have been described to exhibit defects in both rhizoid formation and caulonema differentiation, with double mutants forming significantly smaller gametophytes described as „a dense mat of chloronema[...]" (Tam *et al.*, 2015), with a notable absence of caulonema cells. Like the *PpRSL* genes, the expression of the *PpLRLs* is promoted by auxin, and consequently, double knockouts are auxin-resistant (Tam *et al.*, 2015). Interestingly however, while both set of genes are required to correctly establish caulonema identity, they have been shown to act independently from each other (Jaeger & Moody, 2021; Jang & Dolan, 2011; Tam *et al.*,

2015; Thelander *et al.*, 2018).

Additionally, overexpression of the PIN proteins of *P. patens* has also been reported to result in a comparable phenotype (Viaene *et al.*, 2014). *P. patens* contains four PIN proteins, and overexpression of the long PIN PpPINA or of PpPINB results in the formation of small, dense gametophytes that show a considerable delay in the establishment of caulonema identity, while *pin* knockouts show elongated protonema cells and an earlier transition to caulonema identity. Additionally, overexpression of *A. thaliana* PIN proteins in *P. patens* also disrupts the chloronema – caulonema transition. Is has been suggested that these phenotypes are caused by an increased polar export of cellular auxin into the surrounding medium, thus reducing and delaying the buildup of the local auxin maximum needed for the developmental transition to caulonema identity. In case of *pin* mutants, the inverse effect occurs, with an increased auxin buildup resulting in cell elongation and enhanced caulonema establishment (Viaene *et al.*, 2014).

The common denominator of all three previously described mutants it the auxin signaling pathway, which is the main driver of the chloronema – caulonema transition and as such the most likely pathway affected by the *PpLUG* mutants (Jaeger & Moody, 2021; Kofuji & Hasebe, 2014). While the PpLUGs could theoretically regulate the expression of the *PpRSLs* or *PpLRLs* in addition to, but independently from the auxin pathway, several pieces of evidence refute this hypothesis and indicate a direct regulation of the auxin pathway: For one, while the gametophytes of of both the *pprsl1 pprsl2* and *pplrl1 pplrl2* double mutant lines have been reported to be of around half the size of WT gametophytes (Jang & Dolan, 2011; Tam *et al.*, 2015), the „ball“ gametophytes were considerably smaller than either line, with only around $\sim \frac{1}{50}$ of the size of WT gametophytes (Fig. 14B). Secondly, the callus-like structures of cytokinin-treated „ball“ mutant lines showed a notable absence of leaflet outgrowth, which was present for the WT and the „WT similar“ lines (Fig. 21), and is a hallmark effect of high exogenous cytokinin in *P. patens* (Ashton *et al.*, 1979). Not only are neither PpRSLs nor PpLRLs known to function during gametophore development outside of rhizoids (Jang & Dolan, 2011; Tam *et al.*, 2015; Thelander *et al.*, 2018), but the presence of auxin has been described as a necessary prerequisite for leaflet outgrowth from cytokinin-induced stem cells (Cammarata *et al.*, 2023), indicating a disruption of the auxin pathway in these lines. Thirdly, the „WT similar“ lines exhibited a delay in gametophore induction, a process that is governed by and dependent on both auxin- and cytokinin-mediated regulation (Fig. 15F; Moody, 2022). Finally, the „WT similar“ mutant line #214 showed a significant reduction in sporophyte initiation on top of the gametophores (Fig. 17B). Auxin has previously been linked to both opening of antheridia and archegonia, and to developmental processes in the egg cell and sporophyte (Thelander *et al.*, 2018), which all could explain this striking reduction of reproductive success.

Therefore, it is quite plausible that the PpLUG homologs directly regulate aspects of the auxin signaling pathway, instead of individual gene or gene families. Further, the for-

mation of bistable switches by auxin have been described before: For one, the auxin-mediated regulation via its canonical pathway itself is an example for such a bistable switch (Leyser, 2018). Additionally, there are known feedback loops between auxin signaling and other factors, such as with the SCR – SHORT ROOT protein complex in the root tip stem cell region that regulates asymmetric cell division (Cruz-Ramírez *et al.*, 2012), or with the protein EPIDERMAL PATTERNING FACTOR-LIKE 2, which suppresses cellular auxin response, but without affecting the actual cellular auxin level, to space out the effect of auxin maxima in the leaf margin, thus creating a bistable system for leaf serration regulation (Tameshige *et al.*, 2025). While these examples are precisely coordinated switches instead of resulting from disrupted auxin coordination, they nevertheless set a precedent for the formation of different stable states in conjunction with the auxin signaling pathway.

5.4.3 The LUG – SEU module potentially regulates both auxin homeostasis and components of its canonical signaling pathway

In mosses, auxin is known to to only affect transcription through the TIR1/AFB – Aux/IAA – ARF pathway (Fig. 25A – B; Lavy *et al.*, 2016), while not much is known about the transcriptionally independent rapid auxin response via cAMP signals. In total, there are four *TIR1/AFB* receptors, three *Aux/IAA* genes, and up to 17 *ARFs* present in the *P. patens* genome, alongside three long *PIN* and one short *PIN* genes (Thelander *et al.*, 2018; Viaene *et al.*, 2014). Thus, there are multiple potential points of regulation that could be governed by the LUG – SEU module:

For one, LUG and SEU could regulate the expression of one or more *ARFs*, especially by repressing the expression of class B *ARFs* (Fig. 25C). As these ARFs are generally considered to be transcriptional repressors (Lavy *et al.*, 2016; Mutte *et al.*, 2018; Prigge *et al.*, 2025), loss of LUG and SEU based repression would yield an enhanced class B *ARF* expression and subsequently increased downregulation of their target genes. Indeed, dominant mutants of the class B *ARF*s *PpARFb2* and *PpARFb4* that increase the stability of the proteins against proteasomal degradation have been demonstrated to negatively affect the chloronema – caulonema transition, resulting in phenotypes resembling *pprsl1 pprsl2* and *pplrl1 pplrl2* mutants (Prigge *et al.*, 2025). For the *PpARF2b* stability mutant specifically, the expression of *PpRSL6*, which is also involved in establishing caulonema identity in protonema filaments, is abolished, thus directly linking this repressive class B ARF to the developmental regulation of caulonema establishment (Jaeger & Moody, 2021; Prigge *et al.*, 2025). Therefore, an increased expression of class B ARFs overall can very likely cause a repression of caulonema identity, and depending on the rates of expression and dynamic of auxin-mediated degradation of these ARFs, this repression might also persist under high auxin conditions (Lavy *et al.*, 2016).

Alternatively, the LUG – SEU module could also repress the expression of *Aux/IAA* genes, which in the *pplug* mutant lines would result in a strong transcriptional repression of

class A ARF target genes by the increased presence of Aux/IAA and the subsequent enhanced recruitment of the TPL and TPR corepressors (Fig. 25D; Causier *et al.*, 2012; Lavy *et al.*, 2016; Taylor & Bargmann, 2025). Previous *P. patens* mutants exhibiting mutations in the degron motifs of *Aux/IAA* genes rendering them more stable against proteasomal degradation were reported to both arrest in the chloronema stage, and to be resistant to auxin (Prigge *et al.*, 2010), similar to the *PpARF2b* stability mutants (Prigge *et al.*, 2025). While the general effects of *Aux/IAA* overexpression are scarcely researched, available studies in *A. thaliana* and in *O. sativa* report auxin-insensitive phenotypes for such overexpression lines (Sato & Yamamoto, 2008; Song & Xu, 2013).

Due to the fact the the disruption of caulonema identity specification happens at the first developmental transition in *P. patens* (Jaeger & Moody, 2021), and that all of the *Aux/IAA* degron motif mutants, the *pprsl1 pprsl2* knockouts, and the *pplrl1 pplrl2* mutants were described as completely insensitive to auxin (Jang & Dolan, 2011; Prigge *et al.*, 2010; Tam *et al.*, 2015), it is unfortunately not easily distinguishable whether the LUG – SEU module targets specific class B *ARF*s only, regulates the expression of *Aux/IAA* and therefore indirectly of class A *ARFs*, or functions in both pathways. For *A. thaliana*, AtSEU has been shown to promote expression of the *Aux/IAA* homolog *IAA19*, and *atseu atslk1* double mutants exhibited reduced expression of several components of the auxin signaling pathway in inflorescence meristems and carpels, like the class A ARF *MONOPTEROS*, the class B ARF *ETT*, or the Aux/IAA genes *IAA1* and *IAA17* (Bao *et al.*, 2010; Huai *et al.*, 2018). While it isn't known whether this reduction of expression is due to direct regulation of the genes by AtSEU and AtSLK1, or indirect effects caused by, for example, auxin-based feedback loops, or general disruption of flower development, it nevertheless creates a precedent linking these transcriptional regulators to all levels of the auxin signaling pathway.

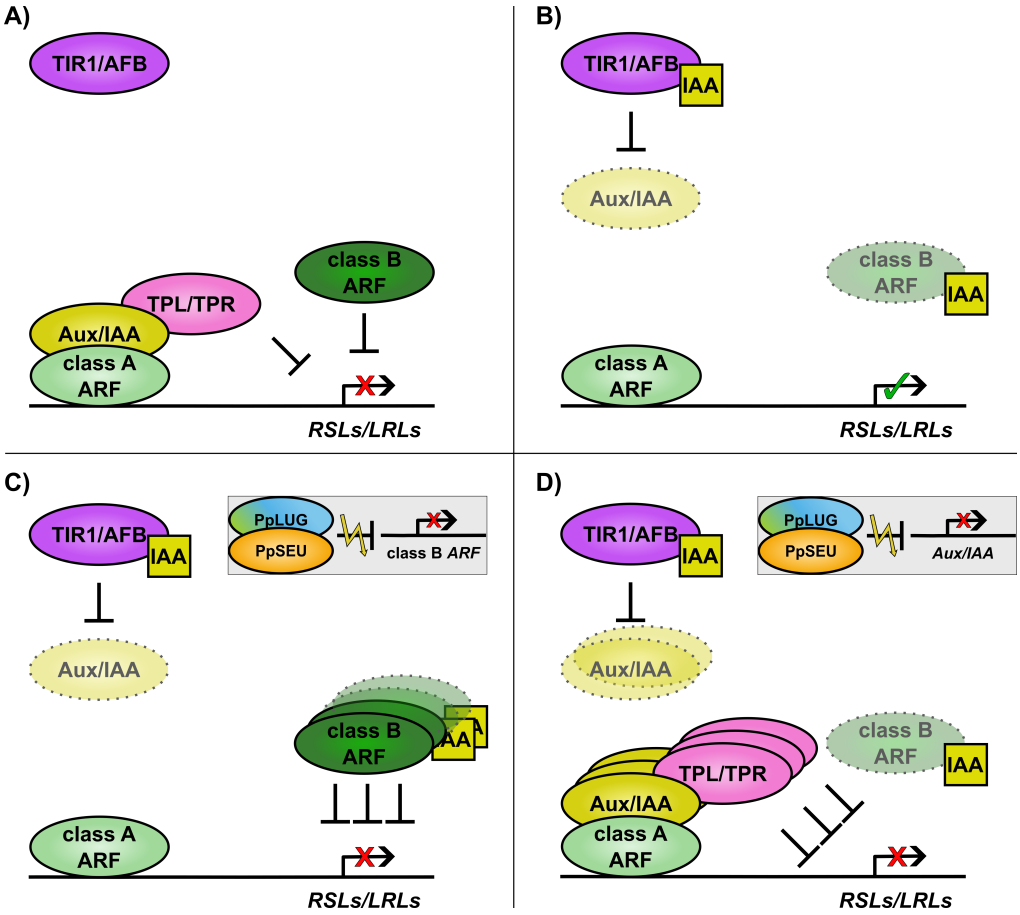


Figure 25: Schematic function of the TIR1/AFB – Aux/IAA – ARF pathway and the role of the LUG – SEU module in the regulation of *PpRSLs* and *PpLRLs*. A) Auxin signaling pathway under low auxin conditions, with Aux/IAA and TPL/TPR repressors inhibiting *PpRSL/PpLRL* expression together with class B ARFs. **B)** High auxin (IAA) conditions cause degradation of Aux/IAA proteins and dissociation of TPL/TPR, resulting in activation of *PpRSL/PpLRL* expression by class A ARFs. **C)** Disruption of the LUG – SEU module mediated repression of class B *ARF*s causes overexpression of class B ARFs and repression of *PpRSL/PpLRL* expression, even under high auxin conditions. **D)** Loss of *Aux/IAA* repression by the LUG – SEU module causes *Aux/IAA* overexpression and repression of class A ARF target genes even under the influence of auxin, due to increased recruitment of TPL/TPR corepressors.

Another possible point of regulation of the auxin signaling pathway by the LUG – SEU complex would be the expression of the *PpPIN* efflux carrier (Fig. 26A – B). A reduction in *PpPIN* inhibition would result in an overproduction of PIN proteins in chloronema cells, which has been demonstrated to delay chloronema – caulonema transition by disruption local auxin maxima due to export of auxin into the surrounding medium (Fig. 26C; Viaene *et al.*, 2014). There are indeed several studies linking LUG and SEU homologs to the regulation of angiosperm PIN proteins: In *A. thaliana*, both AtSEU and AtSLK2 are known to function in *AtPIN1* expression and positioning, with *atseu atslk2* double mutants showing a reduction in cellular auxin accumulation (J. E. Lee *et al.*, 2014). Furthermore, AtLUG has been implicated to regulate auxin homeostasis together with AtSEU by controlling auxin biosynthesis and polar transport (Ståldal *et al.*, 2008). Likewise, the LUG homolog in *A. majus* has been shown to directly or indirectly promote auxin transport (Navarro *et al.*, 2004).

However, this hypothesis is limited by the observation that „ball“ mutants are insensitive

to even high amounts of exogenous auxin (Fig. 19), as it would require the ability of the amassed PIN proteins to reliably deplete cellular auxin under these conditions. Two studies looking at either the dynamic of PIN-mediated efflux or the growth of *A. thaliana* root hairs in various *AtPIN* overexpression lines revealed that, while different PIN proteins do exhibit different auxin efflux rates that are linear to the present auxin concentration, even low amounts of exogenous auxin beneath 50 nM were enough to restore root hair length to control parameters (Ganguly *et al.*, 2010; Janacek *et al.*, 2024). Furthermore, NAA especially has been suggested to passively diffuse into the cell rather than being actively imported, implying that its cellular concentration can only be controlled by polar efflux (Delbarre *et al.*, 1996). Therefore, it is questionable whether an overexpression of *P. patens PINs* would be potent enough to offset the high exogenous NAA concentration. However, it is also important to note that the dynamics of different natural and synthetic auxin variants regarding uptake, transport, and perception remain poorly explored in mosses, as are the kinetics of the different *P. patens* PIN proteins (Thelander *et al.*, 2018; Viaene *et al.*, 2014).

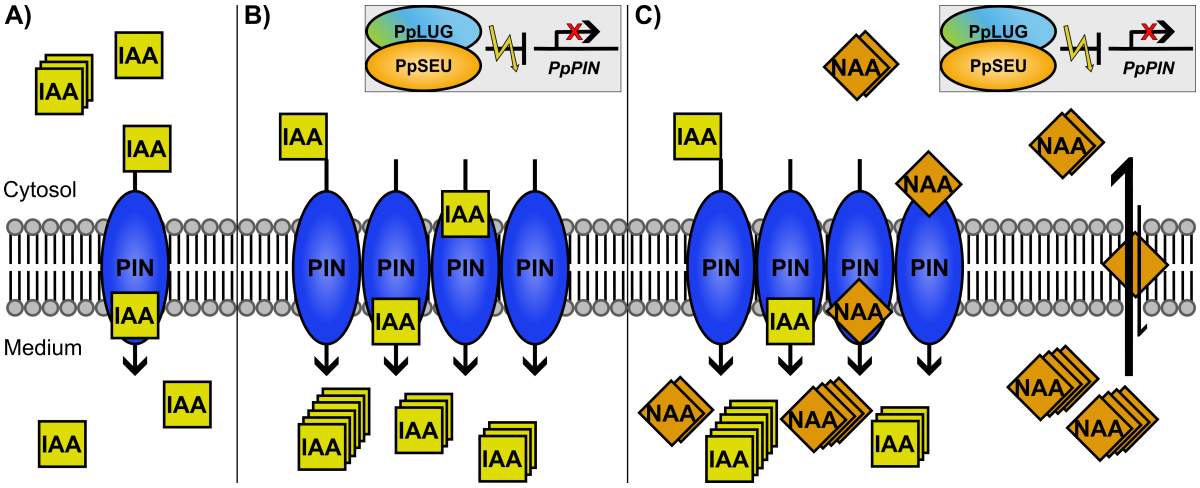


Figure 26: Model for the role of the LUG – SEU module in long PIN efflux carrier function in *P. patens*. **A)** Polar auxin transport through PIN proteins from cytosol to neighbouring cells or the medium. **B)**, **C)** Lack of *PIN* expression repression by the LUG – SEU module causes increased number of PIN proteins and depletion of cellular auxin. **C)** Synthetic exogenous NAA in the medium can diffuse through plasma membrane, following a concentration gradient (Delbarre *et al.*, 1996). Subsequently, PIN proteins export both IAA and NAA, thus possibly maintaining a auxin-depleted cellular state.

Finally, there are two other avenues of regulation of the auxin signaling pathway by the LUG – SEU module, but for both options, no strong supporting evidence is known in neither mosses nor angiosperms: For one, the LUG – SEU complex might also affect the transcriptional activation of *TIR1/AFB* genes, which in turn would reduce the sensitivity of the pathway to auxin (Thelander *et al.*, 2018). Indeed, *ppafp* mutants have been shown to result in loss of caulonema differentiation and auxin response (Prigge *et al.*, 2010). However, no interplay between LUG, SEU, or TIR1/AFB homologs or genes is known across land plants so far. Secondly, AtLUG and AtSEU have been implicated in auxin biosynthesis in angiosperms via regulation of the auxin biosynthesis gene *YUCCA 4*, and of the *SHI/STY*

family of TFs that further promote biosynthesis genes (Bao *et al.*, 2010; Ståldal *et al.*, 2008; Vega-León *et al.*, 2025). While homologs of these families are present in *P. patens* (Thelander *et al.*, 2018), the auxin-insensitivity of the „ball" mutant lines suggest that auxin biosynthesis is unlikely to be affected, or is at least acting in concert with one or more of the other regulatory possibilities discussed above.

In conclusion, there are multiple different aspects of the auxin signaling pathway that could be controlled by the LUG – SEU module. While an exact mechanistical model can't be formulated at this time due to the limited knowledge of auxin regulation by LUG and SEU in angiosperms and in mosses, some substantiated guesses can be made: For one, there is a good chance that the module is responsible for the regulation of multiple facets of the auxin pathway, that together are critically important for auxin sensitivity and the establishment of the caulonema identity during gametophyte development. In *A. thaliana*, AtSEU and AtSLK1 regulate the expression of both class A and B *ARFs*, of multiple *Aux/IAA* homologs, and of auxin synthesis genes (Bao *et al.*, 2010; Huai *et al.*, 2018). Similarly, LUG homologs have been shown to regulate auxin biosynthesis and various auxin signaling factors (Ståldal *et al.*, 2008; Vega-León *et al.*, 2025), emphasizing the multi-factored control of auxin signaling by LUG and SEU in angiosperms, which very well could have originated in the MRCA of land plants during the evolution of the auxin signaling pathway (Hernández-García *et al.*, 2024; Mutte *et al.*, 2018; see section 5.4.4).

Secondly, the changes of chloronema length and width for otherwise auxin-insensitive „ball" mutant lines (Fig. 19C – D) suggest that cell elongation via the rapid auxin response is also present in the bryophyte lineage. Current knowledge about this mechanisms in *P. patens* is scarce (Thelander *et al.*, 2018), but in angiosperms, it is known that the rapid response causes acidification of the apoplast followed by cellular expansion, and acts in concert with, but independently from the transcriptional auxin control of cell growth (Kumar *et al.*, 2025). As the auxin signaling pathway of „ball" mutants is severely disrupted, the observed increase in chloronema length and decrease in chloronema width under high exogenous auxin conditions can best be explained by this separate auxin-mediated response. Consequently, there are multiple different auxin-dependent pathways present in *P. patens*, some of which are independent of the LUG – SEU module and its regulation.

5.4.4 The regulatory functions of LUG and SEU proteins in auxin signaling evolved in the MRCA of land plants

While some streptophyte algae species have been shown to respond to auxin, the emergence of the functional canonical TIR1/AFB – Aux/IAA – ARF signaling pathway has been dated to the MRCA of land plants (Feng *et al.*, 2024; Hernández-García *et al.*, 2024; Mutte *et al.*, 2018). Both the *TIR1/AFB* and the *Aux/IAA* families can be traced across the streptophyte lineage, but the auxin-responding functional module was only gained in the embryophytes, as was the corepression between the Aux/IAAs and the TPL and TPR fami-

lies, as well as the role of the rapid auxin response in cell elongation (Causier *et al.*, 2012; Hernández-García *et al.*, 2024; Mutte *et al.*, 2018; Zeng *et al.*, 2024). In contrast, the origin of the *ARF* family predates the divergence of the Klebsormidiophyceae lineage, with the three classes of *ARFs* sharing a common proto-*ARF* ancestor. Notably, while the class C *ARFs* split from the proto-*ARF* ancestor during the evolution of the streptophyte lineage, the class A and B *ARFs* only diverged in the lineage leading to land plants (Mutte *et al.*, 2018). Consequently, the gene activating regulatory function of class A ARFs was an embryophyte exclusive gain, indicating that transcriptional repression might have been the ancestral function of the *ARF* family (Hernández-García *et al.*, 2024).

Similarly, the origin of the *PIN* family can also be traced back to the early streptophyte evolution, with multiple *Klebsormidium* species exhibiting PIN efflux carriers (Skokan *et al.*, 2019; Viaene *et al.*, 2013). However, these PINs are not polarly localized at cell – cell contact points, but generally distributed at the cell periphery instead. It has been hypothesized that these early PINs might function in quorum sensing and detoxification, and were only later recruited to mediated polar auxin transport (Skokan *et al.*, 2019).

The ability of LUG and SEU homologs to directly or indirectly interact with and regulate the auxin pathway has been demonstrated for several major land plant lineages, such as the angiosperm species *A. thaliana* and *O. sativa* (Bao *et al.*, 2010; J. E. Lee *et al.*, 2014; Pfluger & Zambryski, 2004; Ståldal *et al.*, 2008; Vega-León *et al.*, 2025; Yang *et al.*, 2019), or the moss *P. patens* in this dissertation. Additionally, MpLUG and MpMADS1 have been suggested to influence auxin biosynthesis and signaling in the hornwort *M. polymorpha* as well (Q. Li, 2022), making it highly plausible this interplay emerged alongside the formation of the functional pathway in the MRCA of all land plants. One possibility for the evolution of this connection would be the incorporation, rewiring, and expansion of ancestral regulatory functions LUG and SEU homologs possessed on different auxin pathway components during the transition to land, especially considering that such co-option of available regulatory mechanisms happened multiple times throughout plant evolution (Artur & Kajala, 2021; Bowles *et al.*, 2022). The *LUG*, *SEU*, *ARF*, and *PIN* families all originated before the MRCA of Klebsormidiophyceae and the ZCC-grade (Mutte *et al.*, 2018; Skokan *et al.*, 2019; Wilhelmsson *et al.*, 2017), and thus could have formed ancestral regulatory pathways that got expanded during embryophyta evolution into governing a plethora of different aspect of this novel signaling pathway. This option could also be supported by the known regulation of both class A and class B *ARFs* by the LUG – SEU module in *A. thaliana* (Bao *et al.*, 2010), as their divergence occurred during land plant transition (Mutte *et al.*, 2018). While the regulatory network of such genes tend to separate from each other, there is nevertheless a certain overlap of retained regulatory factors present between them (Maslov *et al.*, 2004).

Another possibility would be the recruitment of the LUG – SEU module to the auxin pathway due to the activity of MADS-box TFs, as several MADS-box proteins are known to

regulate different aspects of auxin signaling in angiosperms, such as AtSEP3, which directly controls multiple *IAA* and *ARF* genes, or the upregulation of the auxin homeostasis controlling GLYCOSIDE HYDROLASE 3 (GH3) family by AtAG (Kaufmann *et al.*, 2009). Similarly, the activity of a GH3 homolog is also controlled by a functional module of OsLUG – OsSEU – OsAP1 – OsSEP3 in *O. sativa* (Yang *et al.*, 2019). However, there are currently no known links between the regulation of the auxin signaling pathway and MADS-box TFs outside of angiosperms, so this connection could also be a lineage-specific acquisition instead of an ancestral function.

Thirdly, there is a known association between protein interaction clusters and biological modules like pathways (Ramos *et al.*, 2024), which could allow for secondary recruitment of LUG and SEU proteins to the regulation of additional members of the pathway after the initial establishment of LUG and / or SEU mediated regulation.

Unfortunately, the exact evolutionary trajectory of the regulation of the different elements the auxin responsive pathway by LUG and SEU homologs can't be elucidated in more detail, due to the sparsity of information beyond the angiosperms and especially land plants. However, the establishment of the pathway control can be unequivocally located in the MRCA of embryophytes, underscoring the critical function of LUG and SEU homologs in a variety of different developmental and regulatory functions throughout the green lineage.

6 Conclusion

The LUG and SEU families represent two evolutionary ancient groups of transcriptional regulators that control a plethora of different gene and regulatory pathways. Their versatility in acting both as a complex or independently, and in conveying both repressing and activating transcriptional control (Gonzalez *et al.*, 2007; Shi *et al.*, 2024; You *et al.*, 2019) signifies their remarkable role as regulatory protein hubs throughout evolution in the green lineage. With their ability to link various different TFs and thus pathways to both histone modifiers and the RNA polymerase II mediator complex (Gonzalez *et al.*, 2007), they became a core aspect of land plant developmental regulation, and possibly also in the streptophyte algae. Their varied interactions with the MADS-box TFs showcase not only the evolutionary conservation of the interaction since the MRCA of the Klebsormidiophyceae and the land plants, but also how their connection to the regulation of sexual development expanded and evolved, rendering the LUG – SEU module a core component in flower development of angiosperms. Similarly, their functions in the auxin signaling pathway, while still only poorly understood, demonstrate well how such regulatory hubs gained relevance in such vast pathways, either by successive recruitment from involved genes into larger regulatory roles, or by co-option of present gene regulatory networks during larger evolutionary shifts, such as the transition from water to land. Thereby, both families represent an interface between sexual reproduction (Azhakanandam *et al.*, 2008; Bao *et al.*, 2010; Conner & Liu, 2000; Franks *et al.*, 2006; Z. Liu *et al.*, 2000; Shi *et al.*, 2024; M. K. Simon *et al.*, 2017; Stahle *et al.*, 2009), phytohormone signaling (J. E. Lee *et al.*, 2014; Pfluger & Zambryski, 2004; Ståldal *et al.*, 2008; Vega-León *et al.*, 2025), stress response (Gong *et al.*, 2016; Shrestha *et al.*, 2014), and more, allowing for a complex control of all aspects of plant development. By taking a variety of different information into account, they help to ensure the proper establishment of various plant organs, and arguably more importantly, the correct formation of the next generation of plants.

7 Outlook

This research into the conservation, interaction, and function of the LUG and SEU families opened up a multitude of further questions about the functional and regulatory origin in the streptophyte algae, their recruitment to various core process of different land plant lineages, and the subsequent evolution of the deeply connected protein interaction networks.

For one, further analyses into the conservation of the interaction between LUG homologs and various histone modifiers or mediator complex proteins could be used to more precisely understand the emergence, conservation, and modulation of this crucial aspect of the LUG – SEU module, especially given the deep conservation of many of these protein families (Alinsug *et al.*, 2009; Nagulapalli *et al.*, 2016). Similarly, analyzing the LUG – SEU interaction in more detail, like the importance of the LisH dimerization motif, the relevance of consistently interacting aa residues and generally of the conserved region of the LUFS domain, or the precise functional roles of the IDRs of SEU for the interaction, could yield a deeper understanding about the interaction as a whole, their evolution, and their conservation.

Furthermore, while functional studies of LUG and SEU homologs in angiosperms and especially *A. thaliana* are well established, they are lacking in many non seed plant lineages. As such, the *P. patens pplug3 pplug4* mutants could be studied in more detail: For one, protein interaction assays like Y2H and CoIP could be used to study whether the proposed disruption of the interaction between PpLUG3ΔLUFS and the PpSEU proteins happened *in vitro*, and to which extend other PpLUG homologs could compensate for this disruption. In combination with an analysis of the general chromatin state and its differences between *P. patens* WT, „WT similar“, and „ball of protonema tissue“ mutant lines, the strength and effects of the disruption on the overall transcriptional regulation could be elucidated. Further, such an analysis could also help in identifying the exact genetic targets of the LUG – SEU module in the auxin pathway. Similarly, an overexpression of PpLUG1 in „ball“ mutant lines could be performed, which together with the other methods could clarify the precise dynamics of the observed bistable state of the auxin signaling pathway in the phenotypes.

Additionally, the creation of both higher order *pplug* and *ppseu* mutants, as well as of inducible *pplug3*ΔLUFS mutant lines might be useful to elucidate further regulatory targets of these homologs (see for example Causier *et al.*, 2023). Especially the latter option would be precious in studying the effects of LUG and SEU homologs on both the caulonema – gametophore transition, and on the development of the gametangia and the sporophyte, given that all of these mechanisms are at least partially dependent on auxin regulation, and are known to be affected in *A. thaliana lug* and *seu* mutants (Conner & Liu, 2000; Grigorova *et al.*, 2011; Z. Liu *et al.*, 2000; Moody, 2022; Shi *et al.*, 2024; Thelander *et al.*, 2018).

Finally, functional studies of *lug* and *seu* mutants in tracheophytes and monilophytes like *C.richardii* might be of interest, as they would allow for a better understanding of the evolution of LUG and SEU mediated regulation in vascular plants, both in context of the auxin signaling pathway, and together with the evolution of the MADS-box TFs throughout land plants, especially considering their blooming relevance in the flower development of angiosperms.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47–97.

Alinsug, M. V., Yu, C.-W., & Wu, K. (2009). Phylogenetic analysis, subcellular localization, and expression patterns of RPD3/HDA1 family histone deacetylases in plants. *BMC Plant Biology*, *9*(1), 37.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403–410.

Alvarez-Buylla, E. R., Benítez, M., Corvera-Poiré, A., Chaos Cador, Á., De Folter, S., Gamboa De Buen, A., Garay-Arroyo, A., García-Ponce, B., Jaimes-Miranda, F., Pérez-Ruiz, R. V., Piñeyro-Nelson, A., & Sánchez-Corrales, Y. E. (2010). Flower Development. *The Arabidopsis Book*, *8*, e0127.

Andreani, J., & Guerois, R. (2014). Evolution of protein interactions: From interactomes to interfaces. *Archives of biochemistry and biophysics*, *554*, 65–75.

Anisimova, M., & Yang, Z. (2007). Multiple Hypothesis Testing to Detect Lineages under Positive Selection that Affects Only a Few Sites. *Molecular Biology and Evolution*, *24*(5), 1219–1228.

Arabidopsis Interactome Mapping Consortium, Dreze, M., Carvunis, A.-R., Charloteaux, B., Galli, M., Pevzner, S. J., Tasan, M., Ahn, Y.-Y., Balumuri, P., Barabási, A.-L., Bautista, V., Braun, P., Byrdsong, D., Chen, H., Chesnut, J. D., Cusick, M. E., Dangl, J. L., De Los Reyes, C., Dricot, A., … Yazaki, J. (2011). Evidence for Network Evolution in an *Arabidopsis* Interactome Map. *Science*, *333*(6042), 601–607.

Artur, M. A. S., & Kajala, K. (2021). Convergent evolution of gene regulatory networks underlying plant adaptations to dry environments. *Plant, Cell & Environment*, *44*(10), 3211–3222.

Ashton, N. W., Grimsley, N. H., & Cove, D. J. (1979). Analysis of gametophytic development in the moss, Physcomitrella patens, using auxin and cytokinin resistant mutants. *Planta*, *144*(5), 427–435.

Azhakanandam, S., Nole-Wilson, S., Bao, F., & Franks, R. G. (2008). *SEUSS* and *AINTEGUMENTA* Mediate Patterning and Ovule Initiation during Gynoecium Medial Domain Development. *Plant Physiology*, *146*(3), 1165–1181.

Balanzà, V., Roig-Villanova, I., Di Marzo, M., Masiero, S., & Colombo, L. (2016). Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. *Development (Cambridge, England)*, *143*(18), 3372–3381.

Bao, F., Azhakanandam, S., & Franks, R. G. (2010). SEUSS and SEUSS-LIKE transcriptional adaptors regulate floral and embryonic development in Arabidopsis. *Plant physiology*, *152*(2), 821–836.

Barabási, A.-L. (2013). Network science. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *371*(1987), 20120375.

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), 6506–6511.

Batada, N. N., Hurst, L. D., & Tyers, M. (2006). Evolutionary and physiological importance of hub proteins. *PLoS computational biology*, *2*(7), e88.

Becker, A., Chen, X., Dresselhaus, T., Gutsche, N., Müller-Schüssele, S. J., Sprunck, S., Theißen, G., Vries, S., & Zachgo, S. (2025). Sexual reproduction in land plants: An evolutionary perspective. *Plant reproduction*, *38*(2), 12.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*(1), 289–300.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *Journal of bacteriology*, *62*(3), 293–300.

Bhattacharya, D., & van Meir, E. G. (2019). A simple genotyping method to detect small CRISPR-Cas9 induced indels by agarose gel electrophoresis. *Scientific reports*, *9*(1), 4437.

Bi, G., Zhao, S., Yao, J., Wang, H., Zhao, M., Sun, Y., Hou, X., Haas, F. B., Varshney, D., Prigge, M., Rensing, S. A., Jiao, Y., Ma, Y., Yan, J., & Dai, J. (2024). Near telomere-to-telomere genome of the model plant Physcomitrium patens. *Nature plants*, *10*(2), 327–343.

Bierenbroodspot, M. J., Pröschold, T., Fürst-Jansen, J. M. R., Vries, S., Irisarri, I., Darienko, T., & Vries, J. (2024). Phylogeny and evolution of streptophyte algae. *Annals of botany*, *134*(3), 385–400.

Bowers, J. E., Chapman, B. A., Rong, J., & Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, *422*(6930), 433–438.

Bowles, A. M. C., Paps, J., & Bechtold, U. (2022). Water-related innovations in land plants evolved by different patterns of gene cooption and novelty. *New Phytologist*, *235*(2), 732–742.

Bowman, J. L. (2022). The origin of a land flora. *Nature plants*, *8*(12), 1352–1369.

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubrmanian, S., Barry, K., Bauer, D., Boehm, C. R., … Schmutz, J. (2017). Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell*, *171*(2), 287–304.e15.

Briginshaw, L. N., Flores-Sandoval, E., Dierschke, T., Alvarez, J. P., & Bowman, J. L. (2022). KANADI promotes thallus differentiation and FR-induced gametangiophore formation in the liverwort Marchantia. *The New phytologist*, *234*(4), 1377–1393.

Brückner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *International Journal of Molecular Sciences*, *10*(6), 2763–2788.

Cammarata, J., Roeder, A. H. K., & Scanlon, M. J. (2023). The ratio of auxin to cytokinin controls leaf development and meristem initiation in *Physcomitrium patens* (R. Napier, Ed.). *Journal of Experimental Botany*, *74*(21), 6541–6550.

Causier, B., Lloyd, J., Stevens, L., & Davies, B. (2012). TOPLESS co-repressor interactions and their evolutionary conservation in plants. *Plant Signaling & Behavior*, *7*(3), 325–328.

Causier, B., McKay, M., Hopes, T., Lloyd, J., Wang, D., Harrison, C. J., & Davies, B. (2023). The TOPLESS corepressor regulates developmental switches in the bryophyte *Physcomitrium patens* that were critical for plant terrestrialisation. *The Plant Journal*, *115*(5), 1331–1344.

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2017). Araport11: A complete reannotation of the Arabidopsis thaliana reference genome. *The Plant journal : for cell and molecular biology*, *89*(4), 789–804.

Clontech. (2008). Yeast Protocols Handbook.

Clontech. (2010). Yeastmaker™ Yeast Transformation System 2 User Manual.

Concordet, J.-P., & Haeussler, M. (2018). CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic acids research*, *46*(W1), W242–W245.

Conner, J., & Liu, Z. (2000). LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(23), 12902–12907.

Conway, S. J., & Di Stilio, V. S. (2020). An ontogenetic framework for functional studies in the model fern Ceratopteris richardii. *Developmental biology*, *457*(1), 20–29.

Crawford, B. C. W., & Yanofsky, M. F. (2008). The formation and function of the female reproductive tract in flowering plants. *Current biology : CB*, *18*(20), R972–8.

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome research*, *14*(6), 1188–1190.

Cruz-Ramírez, A., Díaz-Triviño, S., Blilou, I., Grieneisen, V. A., Sozzani, R., Zamioudis, C., Miskolczi, P., Nieuwland, J., Benjamins, R., Dhonukshe, P., Caballero-Pérez, J., Horvath, B., Long, Y., Mähönen, A. P., Zhang, H., Xu, J., Murray, J. A., Benfey, P. N., Bako, L., … Scheres, B. (2012). A Bistable Circuit Involving SCARECROW-RETINOBLASTOMA Integrates Cues to Inform Asymmetric Stem Cell Division. *Cell*, *150*(5), 1002–1015.

Delbarre, A., Muller, P., Imhoff, V., & Guern, J. (1996). Comparison of mechanisms controlling uptake and accumulation of 2,4-dichlorophenoxy acetic acid, naphthalene-1-acetic acid, and indole-3-acetic acid in suspension-cultured tobacco cells. *Planta*, *198*(4), 532–541.

Di Marzo, M., Babolin, N., Viana, V. E., Oliveira, A. C., Gugi, B., Caporali, E., Herrera-Ubaldo, H., Martínez-Estrada, E., Driouich, A., Folter, S., Colombo, L., & Ezquer, I. (2022). The Genetic Control of SEEDSTICK and LEUNIG-HOMOLOG in Seed and Fruit Development: New Insights into Cell Wall Control. *Plants (Basel, Switzerland)*, *11*(22).

Dong, Y., Chen, S., Cheng, S., Zhou, W., Ma, Q., Chen, Z., Fu, C.-X., Liu, X., Zhao, Y.-p., Soltis, P. S., Wong, G. K.-S., Soltis, D. E., & Xiang, Q.-Y. (2019). Natural selection and repeated patterns of molecular evolution following allopatric divergence. *eLife*, *8*, e45199.

Edwards, K., Johnstone, C., & Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic acids research*, *19*(6), 1349.

Egener, T., Granado, J., Guitton, M.-C., Hohe, A., Holtorf, H., Lucht, J. M., Rensing, S. A., Schlink, K., Schulte, J., Schween, G., Zimmermann, S., Duwenig, E., Rak, B., & Reski, R. (2002). High frequency of phenotypic deviations in Physcomitrella patens plants transformed with a gene-disruption library. *BMC plant biology*, *2*, 6.

Eisenberg, E., & Levanon, E. Y. (2003). Preferential attachment in the protein network evolution. *Physical review letters*, *91*(13), 138701.

Feng, S., Li, N., Chen, H., Liu, Z., Li, C., Zhou, R., Zhang, Y., Cao, R., Ma, X., & Song, X. (2024). Large-scale analysis of the ARF and Aux/IAA gene families in 406 horticultural and other plants. *Molecular Horticulture*, *4*(1), 13.

Fisher, R. A. (1970). *Statistical methods for research workers* (14th ed.). Springer New York.

Folter, S., Immink, R. G. H., Kieffer, M., Parenicová, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., Davies, B., & Angenent, G. C. (2005). Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *The Plant cell*, *17*(5), 1424–1433.

Frank, M. H., & Scanlon, M. J. (2015). Transcriptomic evidence for the evolution of shoot meristem function in sporophyte-dominant land plants through concerted selection of ancestral gametophytic and sporophytic genetic programs. *Molecular biology and evolution*, *32*(2), 355–367.

Franks, R. G., Liu, Z., & Fischer, R. L. (2006). SEUSS and LEUNIG regulate cell proliferation, vascular development and organ polarity in Arabidopsis petals. *Planta*, *224*(4), 801–811.

Franks, R. G., Wang, C., Levin, J. Z., & Liu, Z. (2002). SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. *Development (Cambridge, England)*, *129*(1), 253–263.

Fraser, H. B., Hirsh, A. E., Steinmetz, L. M., Scharfe, C., & Feldman, M. W. (2002). Evolutionary rate in the protein interaction network. *Science (New York, N.Y.)*, *296*(5568), 750–752.

Frerigmann, H., Berger, B., & Gigolashvili, T. (2014). bHLH05 Is an Interaction Partner of MYB51 and a Novel Regulator of Glucosinolate Biosynthesis in Arabidopsis. *Plant Physiology*, *166*(1), 349–369.

Fuchs, J., Jovtchev, G., & Schubert, I. (2008). The chromosomal distribution of histone methylation marks in gymnosperms differs from that of angiosperms. *Chromosome Research*, *16*(6), 891–898.

Ganguly, A., Lee, S. H., Cho, M., Lee, O. R., Yoo, H., & Cho, H.-T. (2010). Differential Auxin-Transporting Activities of PIN-FORMED Proteins in Arabidopsis Root Hair Cells. *Plant Physiology*, *153*(3), 1046–1061.

Garrecht, J., Rössner, C., Sreechakram, V. N. S., Li, Q., Braun, S., Golz, J. F., & Becker, A. (2026). 800 million years of co-evolution in the green plant lineage – the case of LEUNIG and SEUSS transcriptional co-regulators. *Molecular Biology and Evolution*, msag154.

Geng, X., Horst, W. J., Golz, J. F., Lee, J. E., Ding, Z., & Yang, Z.-B. (2017). *LEUNIG_ HOMOLOG* transcriptional co-repressor mediates aluminium sensitivity through *PECTIN METHYLESTERASE 46* -modulated root cell wall pectin methylesterification in Arabidopsis. *The Plant Journal*, *90*(3), 491–504.

Gerlitz, G., Darhin, E., Giorgio, G., Franco, B., & Reiner, O. (2005). Novel Functional Features of the LIS-H Domain: Role in Protein Dimerization, Half-Life and Cellular Localization. *Cell Cycle*, *4*(11), 1632–1640.

Gharib, W. H., & Robinson-Rechavi, M. (2013). The Branch-Site Test of Positive Selection Is Surprisingly Robust but Lacks Power under Synonymous Substitution Saturation and Variation in GC. *Molecular Biology and Evolution*, *30*(7), 1675–1686.

Gillespie, C. S. (2015). Fitting Heavy Tailed Distributions: The poweRlaw Package. *Journal of Statistical Software*, *64*(2).

Goel, A., & Wilkins, M. R. (2012). Dynamic Hubs Show Competitive and Static Hubs Non-Competitive Regulation of Their Interaction Partners (M. Polymenis, Ed.). *PLoS ONE*, *7*(10), e48209.

Goldman, N., & Yang, Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular biology and evolution*, *11*(5), 725–736.

Gong, X., Flores-Vergara, M. A., Hong, J. H., Chu, H., Lim, J., Franks, R. G., Liu, Z., & Xu, J. (2016). SEUSS Integrates Gibberellin Signaling with Transcriptional Inputs from the SHR-SCR-SCL3 Module to Regulate Middle Cortex Formation in the Arabidopsis Root. *Plant Physiology*, *170*(3), 1675–1683.

Gonzalez, D., Bowen, A. J., Carroll, T. S., & Conlan, R. S. (2007). The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator compo-

nents MED14 (SWP) and CDK8 (HEN3) to repress transcription. *Molecular and cellular biology*, *27*(15), 5306–5315.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic acids research*, *40*, D1178–86.

Gramzow, L., & Theissen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome biology*, *11*(6), 214.

Gregis, V., Sessa, A., Colombo, L., & Kater, M. M. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *The Plant cell*, *18*(6), 1373–1382.

Grigoriev, A. (2003). On the number of protein-protein interactions in the yeast proteome. *Nucleic acids research*, *31*(14), 4157–4161.

Grigorova, B., Mara, C., Hollender, C., Sijacic, P., Chen, X., & Liu, Z. (2011). LEUNIG and SEUSS co-repressors regulate *miR172* expression in *Arabidopsis* flowers. *Development*, *138*(12), 2451–2456.

Gu, Z. (2022). Complex heatmap visualization. *iMeta*, *1*(3), e43.

Hall, T. A. (1999). *BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT*.

Han, J.-D. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J. M., Cusick, M. E., Roth, F. P., & Vidal, M. (2004). Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature*, *430*(6995), 88–93.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of molecular biology*, *166*(4), 557–580.

Harris, B. J., Clark, J. W., Schrempf, D., Szöllősi, G. J., Donoghue, P. C. J., Hetherington, A. M., & Williams, T. A. (2022). Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. *Nature ecology & evolution*, *6*(11), 1634–1643.

Hasebe, M., Wen, C. K., Kato, M., & Banks, J. A. (1998). Characterization of MADS homeotic genes in the fern Ceratopteris richardii. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(11), 6222–6227.

Hernández-García, J., Carrillo-Carrasco, V. P., Rienstra, J., Tanaka, K., De Roij, M., Dipp-Álvarez, M., Freire-Ríos, A., Crespo, I., Boer, R., Van Den Berg, W. A. M., Lindhoud, S., & Weijers, D. (2024). Evolutionary origins and functional diversification of Auxin Response Factors. *Nature Communications*, *15*(1), 10909.

Herrera-Ubaldo, H., Campos, S. E., López-Gómez, P., Luna-García, V., Zúñiga-Mayo, V. M., Armas-Caballero, G. E., González-Aguilera, K. L., DeLuna, A., Marsch-Martínez, N., Espinosa-Soto, C., & Folter, S. (2023). The protein-protein interaction landscape of transcription factors during gynoecium development in Arabidopsis. *Molecular plant*, *16*(1), 260–278.

Hickok, L. G. (1977). Cytological relationships between three diploid species of the fern genus Ceratopteris. *Canadian Journal of Botany*, *55*(12), 1660–1667.

Hickok, L. G., & Warne, T. R. (1998). C-Fern Manual. *Carolina Biological Supply Company, Burlington, VT*.

Higo, A., Niwa, M., Yamato, K. T., Yamada, L., Sawada, H., Sakamoto, T., Kurata, T., Shirakawa, M., Endo, M., Shigenobu, S., Yamaguchi, K., Ishizaki, K., Nishihama, R., Kohchi, T., & Araki, T. (2016). Transcriptional Framework of Male Gametogenesis in the Liverwort Marchantia polymorpha L. *Plant & cell physiology*, *57*(2), 325–338.

Hisanaga, T., Fujimoto, S., Cui, Y., Sato, K., Sano, R., Yamaoka, S., Kohchi, T., Berger, F., & Nakajima, K. (2021). Deep evolutionary origin of gamete-directed zygote activation by KNOX/ BELL transcription factors in green plants. *eLife*, *10*.

Hisanaga, T., Wu, S., Schafran, P., Axelsson, E., Akimcheva, S., Dolan, L., Li, F.-W., & Berger, F. (2023). The ancestral chromatin landscape of land plants. *New Phytologist*, *240*(5), 2085–2101.

Hiss, M., Meyberg, R., Westermann, J., Haas, F. B., Schneider, L., Schallenberg-Rüdinger, M., Ullrich, K. K., & Rensing, S. A. (2017). Sexual reproduction, sporophyte development and molecular variation in the model moss Physcomitrella patens: Introducing the ecotype Reute. *The Plant journal : for cell and molecular biology*, *90*(3), 606–620.

Hohe, A., Egener, T., Lucht, J. M., Holtorf, H., Reinhard, C., Schween, G., & Reski, R. (2004). An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens. *Current genetics*, *44*(6), 339–347.

Holehouse, A. S., & Kragelund, B. B. (2024). The molecular basis for cellular function of intrinsically disordered protein regions. *Nature Reviews Molecular Cell Biology*, *25*(3), 187–211.

Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N., Moriyama, T., Ikeuchi, M., Watanabe, M., Wada, H., Kobayashi, K., Saito, M., Masuda, T., Sasaki-Sekimoto, Y., Mashiguchi, K., … Ohta, H. (2014). Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature communications*, *5*, 3978.

Huai, J., Zhang, X., Li, J., Ma, T., Zha, P., Jing, Y., & Lin, R. (2018). SEUSS and PIF4 Coordinately Regulate Light and Temperature Signaling Pathways to Control Plant Growth. *Molecular Plant*, *11*(7), 928–942.

Huang, D., Holtz, W. J., & Maharbiz, M. M. (2012). A genetic bistable switch utilizing nonlinear protein degradation. *Journal of Biological Engineering*, *6*(1), 9.

Huang, Q., Li, W., Fan, R., & Chang, Y. (2014). New MADS-box gene in fern: Cloning and expression analysis of DfMADS1 from Dryopteris fragrans. *PloS one*, *9*(1), e86349.

Invitrogen. (2003). Gateway Technology.

Ispolatov, I. (2005). Binding properties and evolution of homodimers in protein-protein interaction networks. *Nucleic Acids Research*, *33*(11), 3629–3635.

Jaeger, R., & Moody, L. A. (2021). A fundamental developmental transition in Physcomitrium patens is regulated by evolutionarily conserved mechanisms. *Evolution & development*, *23*(3), 123–136.

Janacek, D. P., Kolb, M., Schulz, L., Mergner, J., Kuster, B., Glanc, M., Friml, J., Ten Tusscher, K., Schwechheimer, C., & Hammes, U. Z. (2024). Transport properties of canonical PIN-FORMED proteins from Arabidopsis and the role of the loop domain in auxin transport. *Developmental Cell*, *59*(24), 3259–3271.e4.

Jang, G., & Dolan, L. (2011). Auxin promotes the transition from chloronema to caulonema in moss protonema by positively regulating PpRSL1and PpRSL2 in Physcomitrella patens. *New Phytologist*, *192*(2), 319–327.

Julca, I., Ferrari, C., Flores-Tornero, M., Proost, S., Lindner, A.-C., Hackenberg, D., Steinbachová, L., Michaelidis, C., Gomes Pereira, S., Misra, C. S., Kawashima, T., Borg, M., Berger, F., Goldberg, J., Johnson, M., Honys, D., Twell, D., Sprunck, S., Dresselhaus, T., … Mutwil, M. (2021). Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. *Nature plants*, *7*(8), 1143–1159.

Kadiyala, U., Sprinzak, D., Monk, N. A. M., Taylor, S. E., Verd, B., Sonnen, K. F., Moon, L., Roeder, A. H. K., Perez-Carrasco, R., & Formosa-Jordan, P. (2025). From genes to patterns: Five key dynamical systems concepts to decode developmental regulatory mechanisms. *Development*, *152*(14), dev204617.

Kapoor, B., Kumar, P., Verma, V., Irfan, M., Sharma, R., & Bhargava, B. (2023). How plants conquered land: Evolution of terrestrial adaptation. *Journal of evolutionary biology*, *36*(1), 5–14.

Katoh, K., Misawa, K., Kuma, K.-i., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, *30*(14), 3059–3066.

Kaufmann, K., Melzer, R., & Theißen, G. (2005). MIKC-type MADS-domain proteins: Structural modularity, protein interactions and network evolution in land plants. *Gene*, *347*(2), 183–198.

Kaufmann, K., Muiño, J. M., Jauregui, R., Airoldi, C. A., Smaczniak, C., Krajewski, P., & Angenent, G. C. (2009). Target Genes of the MADS Transcription Factor SEPALLATA3: Integration of Developmental and Hormonal Pathways in the Arabidopsis Flower (D. Weigel, Ed.). *PLoS Biology*, *7*(4), e1000090.

Kerppola, T. K. (2008). Bimolecular Fluorescence Complementation (BiFC) Analysis as a Probe of Protein Interactions in Living Cells. *Annual Review of Biophysics*, *37*(1), 465–487.

Kim, M. H., Cooper, D. R., Oleksy, A., Devedjiev, Y., Derewenda, U., Reiner, O., Otlewski, J., & Derewenda, Z. S. (2004). The Structure of the N-Terminal Domain of the Product

of the Lissencephaly Gene Lis1 and Its Functional Implications. *Structure*, *12*(6), 987–998.

Kim, P. M., Lu, L. J., Xia, Y., & Gerstein, M. B. (2006). Relating Three-Dimensional Structures to Protein Networks Provides Evolutionary Insights. *Science*, *314*(5807), 1938–1941.

Kim, W. K., & Marcotte, E. M. (2008). Age-Dependent Evolution of the Yeast Protein Interaction Network Suggests a Limited Role of Gene Duplication and Divergence (R. Nussinov, Ed.). *PLoS Computational Biology*, *4*(11), e1000232.

Kofuji, R., & Hasebe, M. (2014). Eight types of stem cells in the life cycle of the moss Physcomitrella patens. *Current opinion in plant biology*, *17*, 13–21.

Kolde, R. (2019). *Pheatmap: Pretty Heatmaps*.

Konieczny, L., Roterman-Konieczna, I., & Spólnik, P. (2023). The Structure and Function of Living Organisms. In *Systems Biology* (pp. 1–52). Springer International Publishing.

Krämer, U. (2015). Planting molecular functions in an ecological context with Arabidopsis thaliana. *eLife*, *4*.

Kumar, V., Yadav, S., Heymans, A., & Robert, S. (2025). "Shape of Cell"—An Auxin and Cell Wall Duet. *Physiologia Plantarum*, *177*(3), e70294.

Kunz, C. F., Vries, S., & Vries, J. (2024). Plant terrestrialization: An environmental pull on the evolution of multi-sourced streptophyte phenolics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *379*(1914), 20230358.

Kwantes, M., Liebsch, D., & Verelst, W. (2012). How MIKC* MADS-box genes originated and evidence for their conserved function throughout the evolution of vascular plant gametophytes. *Molecular biology and evolution*, *29*(1), 293–302.

Landberg, K., Pederson, E. R., Viaene, T., Bozorg, B., Friml, J., Jönsson, H., Thelander, M., & Sundberg, E. (2013). The Moss *Physcomitrella patens* Reproductive Organ Development Is Highly Organized, Affected by the Two *SHI/STY* Genes and by the Level of Active Auxin in the *SHI/STY* Expression Domain. *Plant Physiology*, *162*(3), 1406–1419.

Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., Piednoel, M., Gundlach, H., van Bel, M., Meyberg, R., Vives, C., Morata, J., Symeonidi, A., Hiss, M., Muchero, W., Kamisugi, Y., Saleh, O., Blanc, G., Decker, E. L., … Rensing, S. A. (2018). The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. *The Plant journal : for cell and molecular biology*, *93*(3), 515–533.

Lange, M., Orashakova, S., Lange, S., Melzer, R., Theißen, G., Smyth, D. R., & Becker, A. (2013). The seirena B class floral homeotic mutant of California Poppy (Eschscholzia californica) reveals a function of the enigmatic PI motif in the formation of specific multimeric MADS domain protein complexes. *The Plant cell*, *25*(2), 438–453.

Lavy, M., Prigge, M. J., Tao, S., Shain, S., Kuo, A., Kirchsteiger, K., & Estelle, M. (2016). Constitutive auxin response in Physcomitrella reveals complex interactions between Aux/IAA and ARF proteins. *eLife*, *5*, e13325.

Lee, J. E., & Golz, J. F. (2012). Diverse roles of Groucho/Tup1 co-repressors in plant growth and development. *Plant Signaling & Behavior*, *7*(1), 86–92.

Lee, J. E., Lampugnani, E. R., Bacic, A., & Golz, J. F. (2014). SEUSS and SEUSS - LIKE 2 coordinate auxin distribution and KNOXI activity during embryogenesis. *The Plant Journal*, *80*(1), 122–135.

Lee, N., Park, J., Kim, K., & Choi, G. (2015). The Transcriptional Coregulator LEUNIG_HOMOLOG Inhibits Light-Dependent Seed Germination in Arabidopsis. *The Plant Cell*, *27*(8), 2301–2313.

Leyser, O. (2018). Auxin Signaling. *Plant Physiology*, *176*(1), 465–479.

Li, F.-W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M. S., Bennett, T., Blázquez, M. A., Cheng, S., Cuming, A. C., De Vries, J., De Vries, S., Delaux, P.-M., Diop, I. S., Harrison, C. J., Hauser, D., Hernández-García, J., … Szövényi, P. (2020). Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nature Plants*, *6*(3), 259–272.

Li, L.-P., Zhang, B., & Cheng, L. (2022). CPIELA: Computational Prediction of Plant Protein-Protein Interactions by Ensemble Learning Approach From Protein Sequences and Evolutionary Information. *Frontiers in genetics*, *13*, 857839.

Li, Q. (2022, March). *Investigating the function of the LEUNIG regulatory complex in the basal land plant Marchantia polymorpha* [Doctoral dissertation, The University of Melbourne].

Liu, G.-Q., Lian, L., & Wang, W. (2022). The Molecular Phylogeny of Land Plants: Progress and Future Prospects. *Diversity*, *14*(10), 782.

Liu, Z., Franks, R. G., & Klink, V. P. (2000). Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. *The Plant cell*, *12*(10), 1879–1892.

Liu, Z., & Karmarkar, V. (2008). Groucho/Tup1 family co-repressors in plant development. *Trends in plant science*, *13*(3), 137–144.

Lovell, S. C., & Robertson, D. L. (2010). An integrated view of molecular coevolution in protein-protein interactions. *Molecular biology and evolution*, *27*(11), 2567–2575.

Lu, Q., Tang, X., Tian, G., Wang, F., Liu, K., Nguyen, V., Kohalmi, S. E., Keller, W. A., Tsang, E. W. T., Harada, J. J., Rothstein, S. J., & Cui, Y. (2010). Arabidopsis homolog of the yeast TREX-2 mRNA export complex: Components and anchoring nucleoporin. *The Plant journal : for cell and molecular biology*, *61*(2), 259–270.

Luck, K., Kim, D.-K., Lambourne, L., Spirohn, K., Begg, B. E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F. J., Charloteaux, B., Choi, D., Coté, A. G., Daley, M., Deimling, S., Desbuleux, A., Dricot, A., Gebbia, M., Hardy, M. F., Kishore, N., … Calderwood, M. A. (2020). A reference map of the human binary protein interactome. *Nature*, *580*(7803), 402–408.

Lynch, M., & Conery, J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, *290*(5494), 1151–1155.

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., & Butcher, S. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic acids research*, *52*(W1), W521–W525.

Mair, A., Xu, S.-L., Branon, T. C., Ting, A. Y., & Bergmann, D. C. (2019). Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID. *eLife*, *8*, e47864.

Mallett, D. R., Chang, M., Cheng, X., & Bezanilla, M. (2019). Efficient and modular CRISPR-Cas9 vector system for Physcomitrella patens. *Plant Direct*, *3*(9).

Már, M., Nitsenko, K., & Heidarsson, P. O. (2023). Multifunctional Intrinsically Disordered Regions in Transcription Factors. *Chemistry – A European Journal*, *29*(21), e202203369.

Marchant, A., Cisneros, A. F., Dubé, A. K., Gagnon-Arsenault, I., Ascencio, D., Jain, H., Aubé, S., Eberlein, C., Evans-Yamamoto, D., Yachie, N., & Landry, C. R. (2019). The role of structural pleiotropy and regulatory evolution in the retention of heteromers of paralogs. *eLife*, *8*, e46754.

Marchant, D. B., Chen, G., Cai, S., Chen, F., Schafran, P., Jenkins, J., Shu, S., Plott, C., Webber, J., Lovell, J. T., He, G., Sandor, L., Williams, M., Rajasekar, S., Healey, A., Barry, K., Zhang, Y., Sessa, E., Dhakal, R. R., … Chen, Z.-H. (2022). Dynamic genome evolution in a model fern. *Nature Plants*, *8*(9), 1038–1051.

Maslov, S., Sneppen, K., Eriksen, K. A., & Yan, K.-K. (2004). Upstream plasticity and downstream robustness in evolution of molecular networks. *BMC Evolutionary Biology*, *4*(1), 9.

McCourt, R. M., Lewis, L. A., Strother, P. K., Delwiche, C. F., Wickett, N. J., Vries, J., & Bowman, J. L. (2023). Green land: Multiple perspectives on green algal evolution and the earliest land plants. *American journal of botany*, *110*(5), e16175.

Melzer, R., & Theißen, G. (2009). Reconstitution of 'floral quartets' in vitro involving class B and class E floral homeotic proteins. *Nucleic Acids Research*, *37*(8), 2723–2736.

Meng, X., Li, W., Xiang, J., Bedru, H. D., Wang, W., Wu, F.-X., & Li, M. (2022). Temporal-Spatial Analysis of the Essentiality of Hub Proteins in Protein-Protein Interaction Networks. *IEEE Transactions on Network Science and Engineering*, *9*(5), 3504–3514.

Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., Lang, D., Altmann, S., Cyprys, P., Zolg, D. P., Mathieson, T., Bantscheff, M., Hazarika, R. R., Schmidt, T., Dawid, C., … Kuster, B. (2020). Mass-spectrometry-based draft of the Arabidopsis proteome. *Nature*, *579*(7799), 409–414.

Meyberg, R., Perroud, P.-F., Haas, F. B., Schneider, L., Heimerl, T., Renzaglia, K. S., & Rensing, S. A. (2020). Characterisation of evolutionarily conserved key players affecting eu-

karyotic flagellar motility and fertility using a moss model. *The New phytologist*, *227*(2), 440–454.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Haeseler, A., & Lanfear, R. (2020). Corrigendum to: IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular biology and evolution*, *37*(8), 2461.

Moody, L. A. (2022). Unravelling 3D growth in the moss Physcomitrium patens. *Essays in biochemistry*, *66*(6), 769–779.

Musazade, E., Mrisho, I. I., & Feng, X. (2025). Auxin metabolism and signaling: Integrating independent mechanisms and crosstalk in plant abiotic stress responses. *Plant Stress*, *18*, 101034.

Mutte, S. K., Kato, H., Rothfels, C., Melkonian, M., Wong, G. K.-S., & Weijers, D. (2018). Origin and evolution of the nuclear auxin response system. *eLife*, *7*, e33399.

Nagulapalli, M., Maji, S., Dwivedi, N., Dahiya, P., & Thakur, J. K. (2016). Evolution of disorder in Mediator complex and its functional relevance. *Nucleic Acids Research*, *44*(4), 1591–1612.

Navarro, C., Efremova, N., Golz, J. F., Rubiera, R., Kuckenberg, M., Castillo, R., Tietz, O., Saedler, H., & Schwarz-Sommer, Z. (2004). Molecular and genetic interactions between STYLOSA and GRAMINIFOLIA in the control of Antirrhinum vegetative and reproductive development. *Development (Cambridge, England)*, *131*(15), 3649–3659.

Ng, K.-L., & Huang, C.-H. (2004). *Global Topological Study of the Protein-protein Interaction Networks*. arXiv.

Nishiyama, T., Sakayama, H., Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K. K., Haas, F. B., Vanderstraeten, L., Becker, D., Lang, D., Vosolsobě, S., Rombauts, S., Wilhelmsson, P. K. I., Janitza, P., Kern, R., Heyl, A., Rümpler, F., Villalobos, L. I. A. C., Clay, J. M., … Rensing, S. A. (2018). The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. *Cell*, *174*(2), 448–464.e24.

Nooren, I. M. (2003). Diversity of protein-protein interactions. *The EMBO Journal*, *22*(14), 3486–3492.

Oya, S., Takahashi, M., Takashima, K., Kakutani, T., & Inagaki, S. (2022). Transcription-coupled and epigenome-encoded mechanisms direct H3K4 methylation. *Nature Communications*, *13*(1), 4521.

Pandey, R. (2002). Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research*, *30*(23), 5036–5055.

Parvathy, J., Yazhini, A., Srinivasan, N., & Sowdhamini, R. (2024). Interfacial residues in protein-protein complexes are in the eyes of the beholder. *Proteins*, *92*(4), 509–528.

Pfannebecker, K. C., Lange, M., Rupp, O., & Becker, A. (2017). An Evolutionary Framework for Carpel Developmental Control Genes. *Molecular biology and evolution*, *34*(2), 330–348.

Pfluger, J., & Zambryski, P. (2004). The role of SEUSS in auxin response and floral organ patterning. *Development (Cambridge, England)*, *131*(19), 4697–4707.

Posit team. (2024). *RStudio: Integrated Development Environment for R*. Boston, MA.

Povolotskaya, I. S., & Kondrashov, F. A. (2010). Sequence space and the ongoing expansion of the protein universe. *Nature*, *465*(7300), 922–926.

Prigge, M. J., Lavy, M., Ashton, N. W., & Estelle, M. (2010). Physcomitrella patens Auxin-Resistant Mutants Affect Conserved Elements of an Auxin-Signaling Pathway. *Current Biology*, *20*(21), 1907–1912.

Prigge, M. J., Morffy, N., De Neve, A., Szutu, W., Abraham-Juárez, M. J., McAllister, T., Jones, H., Johnson, K., Do, N., Lavy, M., Hake, S., Strader, L. C., Estelle, M., & Richardson, A. E. (2025). Comparative mutant analyses reveal a novel mechanism of ARF regulation in land plants. *Nature Plants*, *11*(4), 821–835.

Priyam, A., Woodcroft, B. J., Rai, V., Moghul, I., Munagala, A., Ter, F., Chowdhary, H., Pieniak, I., Maynard, L. J., Gibbins, M. A., Moon, H., Davis-Richardson, A., Uludag, M., Watson-Haigh, N. S., Challis, R., Nakamura, H., Favreau, E., Gómez, E. A., Pluskal, T., … Wurm, Y. (2019). Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databas-es. *Molecular biology and evolution*, *36*(12), 2922–2924.

Przytycka, T. M., Singh, M., & Slonim, D. K. (2010). Toward the dynamic interactome: It's about time. *Briefings in Bioinformatics*, *11*(1), 15–29.

Qiu, Y., & Köhler, C. (2022). Endosperm Evolution by Duplicated and Neofunctionalized Type I MADS-Box Transcription Factors. *Molecular biology and evolution*, *39*(1).

Qiu, Y., & Köhler, C. (2026). The Other Giant: Functional Significance and Divergence of Type I MADS-Box Genes in the Evolution of Land Plants. *Journal of Experimental Botany*, erag134.

Qiu, Y., Li, Z., Walther, D., & Köhler, C. (2023). Updated Phylogeny and Protein Structure Predictions Revise the Hypothesis on the Origin of MADS-box Transcription Factors in Land Plants. *Molecular biology and evolution*, *40*(9).

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Ramos, R. H., Oliveira Lage Ferreira, C., & Simao, A. (2024). Human protein-protein interaction networks: A topological comparison review. *Heliyon*, *10*(5), e27278.

Rensing, S. A., Goffinet, B., Meyberg, R., Wu, S.-Z., & Bezanilla, M. (2020). The Moss Physcomitrium (Physcomitrella) patens: A Model Organism for Non-Seed Plants. *The Plant cell*, *32*(5), 1361–1376.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.-F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S.-I., … Boore,

J. L. (2008). The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science (New York, N.Y.)*, *319*(5859), 64–69.

Reski, R., & Abel, W. O. (1985). Induction of budding on chloronemata and caulonemata of the moss, Physcomitrella patens, using isopentenyladenine. *Planta*, *165*(3), 354–358.

Rodriguez-Rivas, J., Marsili, S., Juan, D., & Valencia, A. (2016). Conservation of coevolving protein interfaces bridges prokaryote-eukaryote homologies in the twilight zone. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(52), 15018–15023.

Roessner, C., Griep, S., & Becker, A. (2024). A land plant phylogenetic framework for GLABROUS INFLORESCENCE STEMS (GIS), SUPERMAN, JAGGED and allies plus their TOPLESS co-repressor. *Molecular phylogenetics and evolution*, *201*, 108195.

Rümpler, F., Tessari, C., Gramzow, L., Gafert, C., Blohs, M., & Theißen, G. (2023). The Origin of Floral Quartet Formation-Ancient Exon Duplications Shaped the Evolution of MIKC-type MADS-domain Transcription Factor Interactions. *Molecular biology and evolution*, *40*(5).

Saez-Aguayo, S., Ralet, M.-C., Berger, A., Botran, L., Ropartz, D., Marion-Poll, A., & North, H. M. (2013). PECTIN METHYLESTERASE INHIBITOR6 promotes Arabidopsis mucilage release by limiting methylesterification of homogalacturonan in seed coat epidermal cells. *The Plant cell*, *25*(1), 308–323.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning : A laboratory manual. *Cold Spring Harbor Laboratory Press*.

Sato, A., & Yamamoto, K. T. (2008). Overexpression of the non-canonical *Aux/IAA* genes causes auxin-related aberrant phenotypes in Arabidopsis. *Physiologia Plantarum*, *133*(2), 397–405.

Sauret-Güeto, S., Frangedakis, E., Silvestri, L., Rebmann, M., Tomaselli, M., Markel, K., Delmans, M., West, A., Patron, N. J., & Haseloff, J. (2020). Systematic Tools for Reprogramming Plant Gene Expression in a Simple Model, Marchantia polymorpha. *ACS synthetic biology*, *9*(4), 864–882.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature methods*, *9*(7), 676–682.

Schneider, T. D., & Stephens, R. M. (1990). Sequence logos: A new way to display consensus sequences. *Nucleic acids research*, *18*(20), 6097–6100.

Schrödinger, L. L. (2010). *The PyMOL Molecular Graphics System, Version 3.0*.

Schween, G., Hohe, A., Koprivova, A., & Reski, R. (2003). Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, Physcomitrella patens. *Journal of plant physiology*, *160*(2), 209–212.

Serrano-Pérez, E., Romero-Losada, A. B., Morales-Pineda, M., García-Gómez, M. E., Couso, I., García-González, M., & Romero-Campero, F. J. (2022). Transcriptomic and Metabolomic Response to High Light in the Charophyte Alga Klebsormidium nitens. *Frontiers in plant science*, *13*, 855243.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*(11), 2498–2504.

Shi, L., Li, C., Lv, G., Li, X., Feng, W., Bi, Y., Wang, W., Wang, Y., Zhu, L., Tang, W., & Fu, Y. (2024). The adaptor protein ECAP, the corepressor LEUNIG, and the transcription factor BEH3 interact and regulate microsporocyte generation in Arabidopsis. *The Plant cell*, *36*(7), 2531–2549.

Shimamura, M. (2016). Marchantia polymorpha: Taxonomy, Phylogeny and Morphology of a Model System. *Plant & cell physiology*, *57*(2), 230–256.

Shrestha, B., Guragain, B., & Sridhar, V. V. (2014). Involvement of co-repressor LUH and the adapter proteins SLK1 and SLK2 in the regulation of abiotic stress response genes in Arabidopsis. *BMC plant biology*, *14*, 54.

Simon, M. K., Skinner, D. J., Gallagher, T. L., & Gasser, C. S. (2017). Integument Development in Arabidopsis Depends on Interaction of YABBY Protein INNER NO OUTER with Coactivators and Corepressors. *Genetics*, *207*(4), 1489–1500.

Simon, S., & Petrášek, J. (2011). Why plants need more than one type of auxin. *Plant Science*, *180*(3), 454–460.

Simonini, S., Deb, J., Moubayidin, L., Stephenson, P., Valluru, M., Freire-Rios, A., Sorefan, K., Weijers, D., Friml, J., & Østergaard, L. (2016). A noncanonical auxin-sensing mechanism is required for organ morphogenesis in *Arabidopsis*. *Genes & Development*, *30*(20), 2286–2296.

Sitaraman, J., Bui, M., & Liu, Z. (2008). LEUNIG_HOMOLOG and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. *Plant physiology*, *147*(2), 672–681.

Skern-Mauritzen, R., & Mikkelsen, T. N. (2021). The information continuum model of evolution. *Biosystems*, *209*, 104510.

Skokan, R., Medvecká, E., Viaene, T., Vosolsobě, S., Zwiewka, M., Müller, K., Skůpa, P., Karady, M., Zhang, Y., Janacek, D. P., Hammes, U. Z., Ljung, K., Nodzyński, T., Petrášek, J., & Friml, J. (2019). PIN-driven auxin transport emerged early in streptophyte evolution. *Nature Plants*, *5*(11), 1114–1119.

Smith, T., & Waterman, M. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology*, *147*(1), 195–197.

Soltis, D. E., Visger, C. J., Marchant, D. B., & Soltis, P. S. (2016). Polyploidy: Pitfalls and paths to a paradigm. *American Journal of Botany*, *103*(7), 1146–1166.

Song, Y., & Xu, Z.-F. (2013). Ectopic Overexpression of an AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA) Gene OsIAA4 in Rice Induces Morphological Changes and Reduces Re-

sponsiveness to Auxin. *International Journal of Molecular Sciences*, *14*(7), 13645–13656.

Sreedasyam, A., Plott, C., Hossain, M. S., Lovell, J. T., Grimwood, J., Jenkins, J. W., Daum, C., Barry, K., Carlson, J., Shu, S., Phillips, J., Amirebrahimi, M., Zane, M., Wang, M., Goodstein, D., Haas, F. B., Hiss, M., Perroud, P.-F., Jawdy, S. S., … Schmutz, J. (2023). JGI Plant Gene Atlas: An updateable transcriptome resource to improve functional gene descriptions across the plant kingdom. *Nucleic acids research*, *51*(16), 8383–8401.

Sridhar, V. V., Surendrarao, A., Gonzalez, D., Conlan, R. S., & Liu, Z. (2004). Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(31), 11494–11499.

Sridhar, V. V., Surendrarao, A., & Liu, Z. (2006). APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. *Development (Cambridge, England)*, *133*(16), 3159–3166.

Stahle, M. I., Kuehlich, J., Staron, L., Arnim, A. G., & Golz, J. F. (2009). YABBYs and the transcriptional corepressors LEUNIG and LEUNIG_HOMOLOG maintain leaf polarity and meristem activity in Arabidopsis. *The Plant cell*, *21*(10), 3105–3118.

Ståldal, V., Sohlberg, J. J., Eklund, D. M., Ljung, K., & Sundberg, E. (2008). Auxin can act independently of *CRC* , *LUG* , *SEU* , *SPT* and *STY1* in style development but not apical-basal patterning of the *Arabidopsis* gynoecium. *New Phytologist*, *180*(4), 798–808.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksöz, E., Droege, A., Krobitsch, S., … Wanker, E. E. (2005). A human protein-protein interaction network: A resource for annotating the proteome. *Cell*, *122*(6), 957–968.

Striebinger, H., Koegl, M., & Bailer, S. M. (2013). A high-throughput yeast two-hybrid protocol to determine virus-host protein interactions. *Methods in molecular biology (Clifton, N.J.)*, *1064*, 1–15.

Student. (1908). The probable error of a mean. *Biometrika*, 1–25.

Sun, M. G. F., & Kim, P. M. (2011). Evolution of biological interaction networks: From models to real data. *Genome biology*, *12*(12), 235.

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic acids research*, *34*, W609–12.

Tam, T. H. Y., Catarino, B., & Dolan, L. (2015). Conserved regulatory mechanism controls the development of cells with rooting functions in land plants. *Proceedings of the National Academy of Sciences*, *112*(29).

Tameshige, T., Tsuchida, T., Matsushita, Y., Doll, Y., Maruyama, K., Agui, T., Aida, M., Kasahara, H., Torii, K. U., Uchida, N., Fujimoto, K., & Ikeuchi, M. (2025). Mutual inhibition between EPFL2 and auxin extends the intervals of periodic leaf morphogenesis. *Nature Communications*, *16*(1), 9753.

Tanabe, Y., Hasebe, M., Sekimoto, H., Nishiyama, T., Kitani, M., Henschel, K., Münster, T., Theissen, G., Nozaki, H., & Ito, M. (2005). Characterization of MADS-box genes in charophycean green algae and its implication for the evolution of MADS-box genes. *Proceedings of the National Academy of Sciences*, *102*(7), 2436–2441.

Tanizawa, Y., Mochizuki, T., Yagura, M., Sakamoto, M., Fujisawa, T., Kawamura, S., Shimokawa, E., Yamaoka, S., Nishihama, R., Bowman, J. L., Berger, F., Yamato, K., Kohchi, T., & Nakamura, Y. (2025). *MarpolBase: Genome database for Marchantia polymorpha featuring high quality reference genome sequences*.

Taylor, J. S., & Bargmann, B. O. R. (2025). Transcriptional Tuning: How Auxin Strikes Unique Chords in Gene Regulation. *Physiologia Plantarum*, *177*(3), e70229.

Thangavel, G., & Nayar, S. (2018). A Survey of MIKC Type MADS-Box Genes in Non-seed Plants: Algae, Bryophytes, Lycophytes and Ferns. *Frontiers in plant science*, *9*, 510.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815.

The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Adesina, A., Ahmad, S., Bowler-Barnett, E. H., Bye-A-Jee, H., Carpentier, D., Denny, P., Fan, J., Garmiri, P., Gonzales, L. J. D. C., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Joshi, V., … Zhang, J. (2025). UniProt: The Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*, *53*(D1), D609–D617.

Thelander, M., Landberg, K., & Sundberg, E. (2018). Auxin-mediated developmental control in the moss Physcomitrella patens. *Journal of Experimental Botany*, *69*(2), 277–290.

Thelander, M., Olsson, T., & Ronne, H. (2005). Effect of the energy supply on filamentous growth and development in Physcomitrella patens. *Journal of experimental botany*, *56*(412), 653–662.

Thompson, J. N. (1994). *The coevolutionary process*. University of Chicago Press.

van de Peer, Y., Maere, S., & Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature reviews. Genetics*, *10*(10), 725–732.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., & Babu, M. M. (2014). Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews*, *114*(13), 6589–6631.

van Meyel, D. J., Thomas, J. B., & Agulnick, A. D. (2003). Ssdp proteins bind to LIM-interacting co-factors and regulate the activity of LIM-homeodomain protein complexes in vivo. *Development (Cambridge, England)*, *130*(9), 1915–1925.

Vanneste, K., Van De Peer, Y., & Maere, S. (2013). Inference of Genome Duplications from Age Distributions Revisited. *Molecular Biology and Evolution*, *30*(1), 177–190.

Vega-León, R., Muino, J. M., Zhu, T., Gluza, J. A., Pimentel, M., Rehbein, K., Steiner, S., Rezzolla, E., Golz, J. F., Chen, D., Kaufmann, K., & Smaczniak, C. (2025). Rewiring of LEUNIG_HOMOLOG interaction networks marks regulatory shifts from meristem to organ growth in Arabidopsis flowers. *The Plant Journal*, *124*(4), e70585.

Viaene, T., Delwiche, C. F., Rensing, S. A., & Friml, J. (2013). Origin and evolution of PIN auxin transporters in the green lineage. *Trends in Plant Science*, *18*(1), 5–10.

Viaene, T., Landberg, K., Thelander, M., Medvecka, E., Pederson, E., Feraru, E., Cooper, E. D., Karimi, M., Delwiche, C. F., Ljung, K., Geisler, M., Sundberg, E., & Friml, J. (2014). Directional Auxin Transport Mechanisms in Early Diverging Land Plants. *Current Biology*, *24*(23), 2786–2791.

Vries, J., & Archibald, J. M. (2018). Plant evolution: Landmarks on the path to terrestrial life. *The New phytologist*, *217*(4), 1428–1434.

Wagih, O., & Parts, L. (2014). Gitter: A Robust and Accurate Method for Quantification of Colony Sizes From Plate Images. *G3 Genes|Genomes|Genetics*, *4*(3), 547–552.

Wagner, G. P. (2010). THE MEASUREMENT THEORY OF FITNESS. *Evolution*.

Walker, M., Tehseen, M., Doblin, M. S., Pettolino, F. A., Wilson, S. M., Bacic, A., & Golz, J. F. (2011). The transcriptional regulator LEUNIG_HOMOLOG regulates mucilage release from the Arabidopsis testa. *Plant physiology*, *156*(1), 46–60.

Walter, M., Chaban, C., Schütze, K., Batistic, O., Weckermann, K., Näke, C., Blazevic, D., Grefen, C., Schumacher, K., Oecking, C., Harter, K., & Kudla, J. (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *The Plant journal : for cell and molecular biology*, *40*(3), 428–438.

Wang, B., Zhang, H., Huai, J., Peng, F., Wu, J., Lin, R., & Fang, X. (2022). Condensation of SEUSS promotes hyperosmotic stress tolerance in Arabidopsis. *Nature chemical biology*, *18*(12), 1361–1369.

Wang, H., Wang, Z., Tang, Q., Yan, X.-X., & Xu, W. (2019). Crystal structure of the LUFS domain of human single-stranded DNA binding Protein 2 (SSBP2). *Protein Science*, *28*(4), 788–793.

Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2023). The conserved domain database in 2023. *Nucleic acids research*, *51*(D1), D384–D388.

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*.

Wilhelmsson, P. K. I., Mühlich, C., Ullrich, K. K., & Rensing, S. A. (2017). Comprehensive Genome-Wide Classification Reveals That Many Plant-Specific Transcription Factors Evolved in Streptophyte Algae. *Genome Biology and Evolution*, *9*(12), 3384–3397.

Xiao, L., & Xia, K. (2025). Functions of Intrinsically Disordered Regions. *Biology*, *14*(7), 810.

Xu, C., & Min, J. (2011). Structure and function of WD40 domain proteins. *Protein & Cell*, *2*(3), 202–214.

Yang, C., Liu, X., Li, D., Zhu, X., Wei, Z., Feng, Z., Zhang, L., He, J., Mou, C., Jiang, L., & Wan, J. (2019). OsLUGL is involved in the regulating auxin level and OsARFs expression in rice (Oryza sativa L.) *Plant Science*, *288*, 110239.

You, Y., Zhai, Q., An, C., & Li, C. (2019). LEUNIG_HOMOLOG Mediates MYC2-Dependent Transcriptional Activation in Cooperation with the Coactivators HAC1 and MED25. *The Plant cell*, *31*(9), 2187–2205.

Zeng, H. Y., Deng, S., Jin, C., Shang, Z., Chang, L., Wang, J., Han, X., Wang, A., Jin, D., Wang, Y., He, H., Li, L., Deng, X. W., & Wei, N. (2024). Origin and evolution of auxin-mediated acid growth. *Proceedings of the National Academy of Sciences*, *121*(51), e2412493121.

Zhai, H., Zhang, X., You, Y., Lin, L., Zhou, W., & Li, C. (2020). SEUSS integrates transcriptional and epigenetic control of root stem cell organizer specification. *The EMBO journal*, *39*(20), e105047.

Zhang, F., Wang, H., Kalve, S., Wolabu, T. W., Nakashima, J., Golz, J. F., & Tadege, M. (2019). Control of leaf blade outgrowth and floral organ development by LEUNIG , ANGUSTIFOLIA 3 and WOX transcriptional regulators. *New Phytologist*, *223*(4), 2024–2038.

Zhang, L., Ma, F., Duan, G., Ju, Y., Yu, T., Zhang, Q., & Sodmergen. (2024). MIKC*-type MADS transcription factors control JINGUBANG expression and the degree of pollen dormancy in Arabidopsis. *Plant physiology*, *197*(1).

Zhang, M.-X., Zhu, S.-S., Xu, Y.-C., Guo, Y.-L., Yang, W.-C., & Li, H.-J. (2020). Transcriptional repression specifies the central cell for double fertilization. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(11), 6231–6236.

Zhang, Z. (2022). KaKs_Calculator 3.0: Calculating Selective Pressure on Coding and Non-coding Sequences. *Genomics, proteomics & bioinformatics*, *20*(3), 536–540.

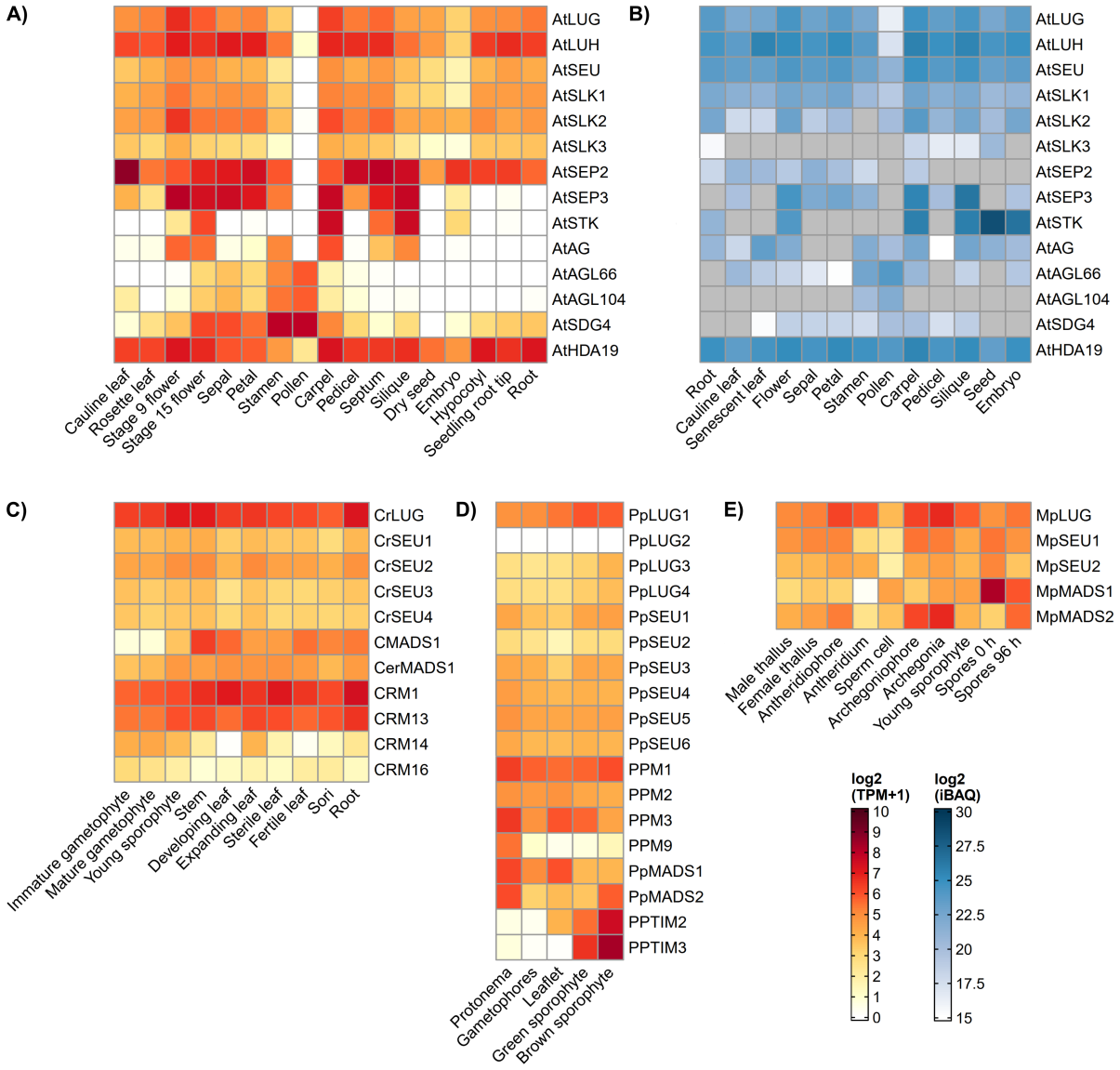
Zhang, Z., Zou, W., Lin, P., Wang, Z., Chen, Y., Yang, X., Zhao, W., Zhang, Y., Wang, D., Que, Y., & Wu, Q. (2024). Evolution and Function of MADS-Box Transcription Factors in Plants. *International Journal of Molecular Sciences*, *25*(24), 13278.

Zobell, O., Faigl, W., Saedler, H., & Münster, T. (2010). MIKC* MADS-box proteins: Conserved regulators of the gametophytic generation of land plants. *Molecular biology and evolution*, *27*(5), 1201–1211.

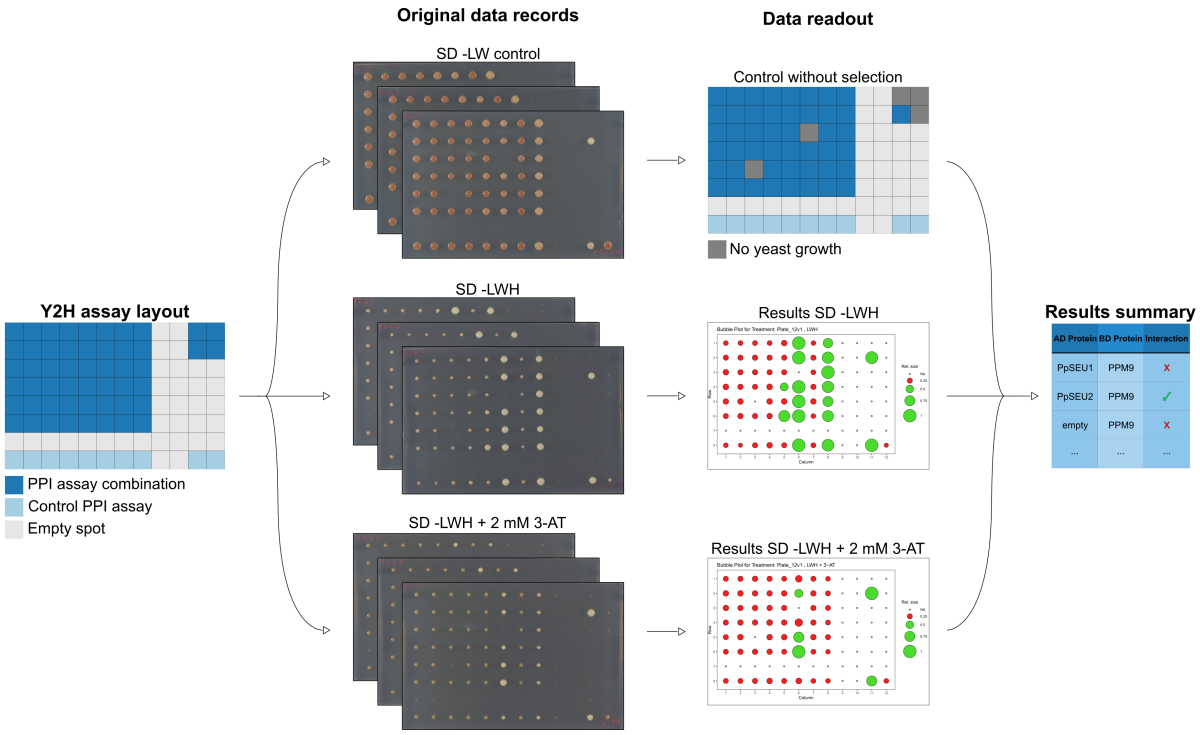
Zubarev, R. A. (2013). The challenge of the proteome dynamic range and its implications for in-depth proteomics. *PROTEOMICS*, *13*(5), 723–726.

Supplementary Material

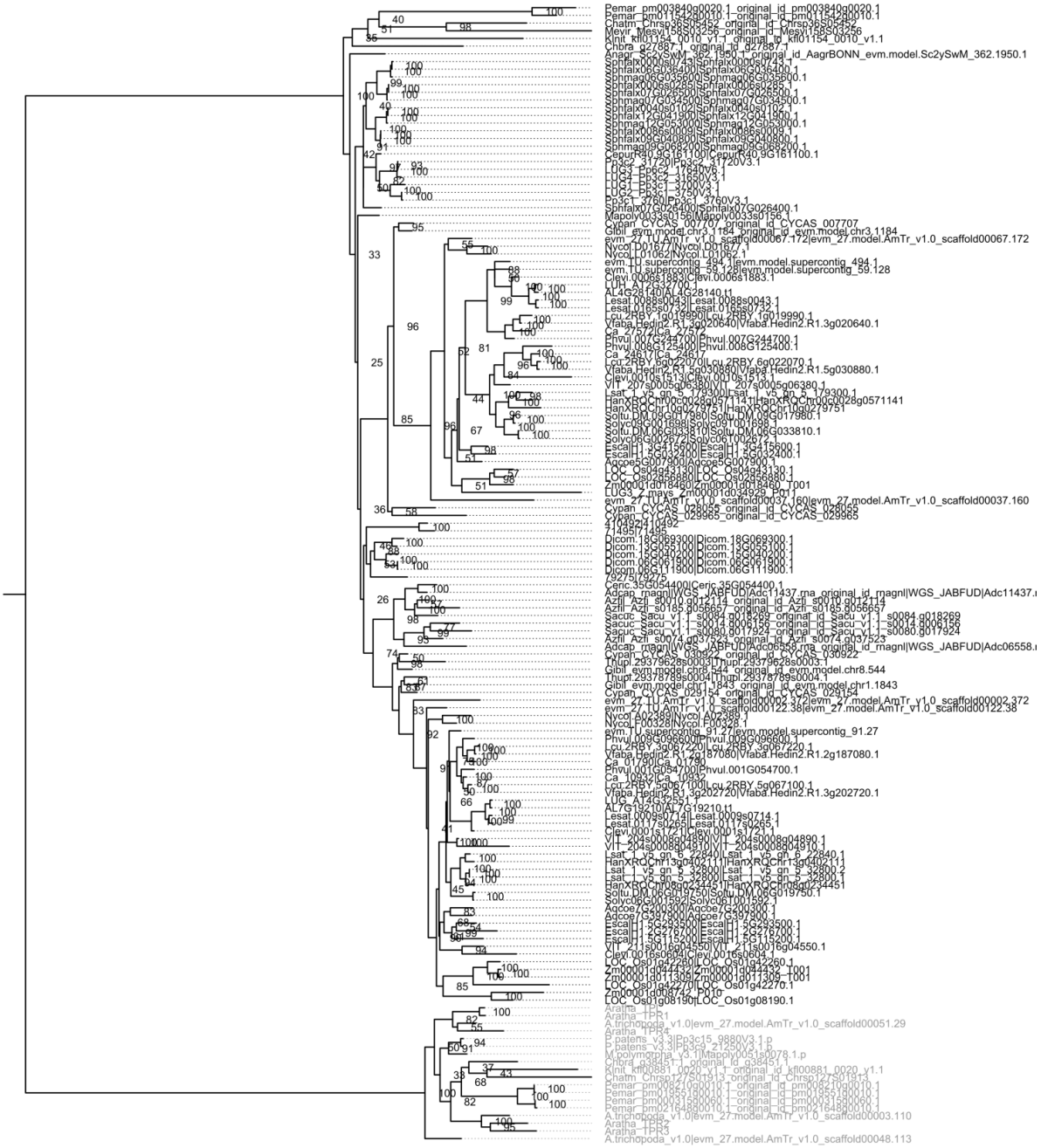
Supplementary Figures and Tables



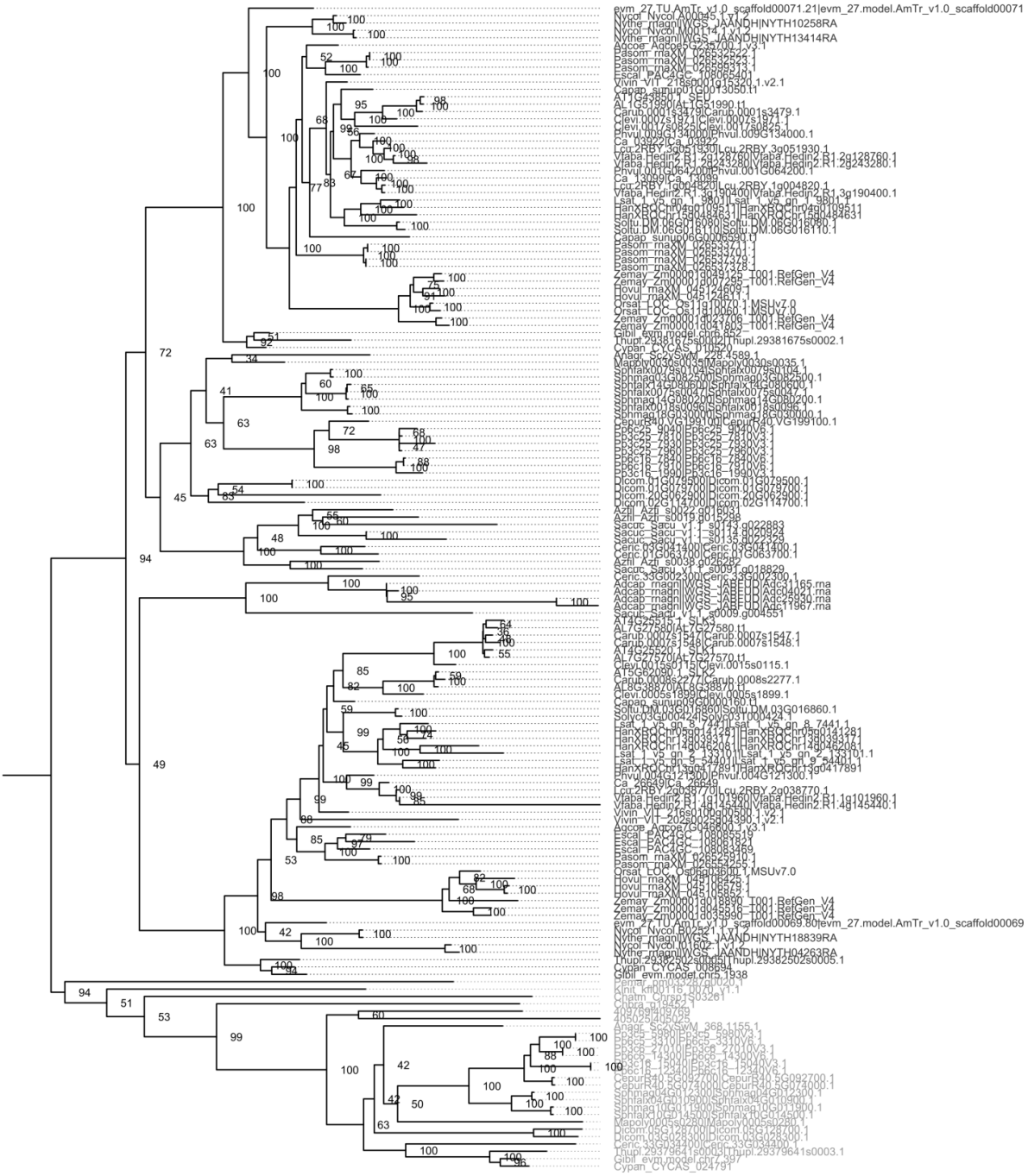
Supplementary Figure 1: Comparative gene expression and proteome analysis of *LUG, SEU*, and MADS-box homologs. Gene expression and protein abundance of *LUG*, *SEU*, and MADS-box homologs across different developmental tissues and stages, as $\log_2$(TPM + 1) or $\log_2$(iBAQ), with identical color scales each. **A)** Gene expression and **B)** protein abundance values for *A. thaliana* Col-0 genes and proteins, respectively (Mergner *et al.*, 2020). The histone remodelers *AtSDG4* and *AtHDA19* were added as well. **C)** Gene expression for *C. richardii* strain Hn-n (D. B. Marchant *et al.*, 2022); for **D)** *P. patens* Gransden 2004 (Lang *et al.*, 2018; Sreedasyam *et al.*, 2023); and for *M. polymorpha* 15 d male and female thallus (Melbourne strain) (Briginshaw *et al.*, 2022); Tak-1 sperm cells (Julca *et al.*, 2021), antheridiophores, and antheridia (Higo *et al.*, 2016); Tak-2 archegoniophores (Higo *et al.*, 2016) and archegonia (Hisanaga *et al.*, 2021); BC3 x Tak-1 13 d old sporophytes (Frank & Scanlon, 2015); and Cam1 x Cam2 spores (Bowman *et al.*, 2017).



Supplementary Figure 2: Automated analysis of Y2H assays. Yeast colonies were grown on basic synthetic dropout SD -LW medium as control, on SD -LWH medium, and on -LWH medium containing 2 mM 3-AT as competitive inhibitor. Each step was performed with three biological replicates and in a standardized layout. Imaging of incubated plates was done from a fixed distance, followed by quantification of colony growth. Relative colony sizes were calculated by dividing colony size on -LWH medium by colony size on corresponding -LW control medium. Relative sizes ≥ 0.5 were considered interacting proteins, while empty spots were discarded prior to analysis. Results were compiled across all three replicates.



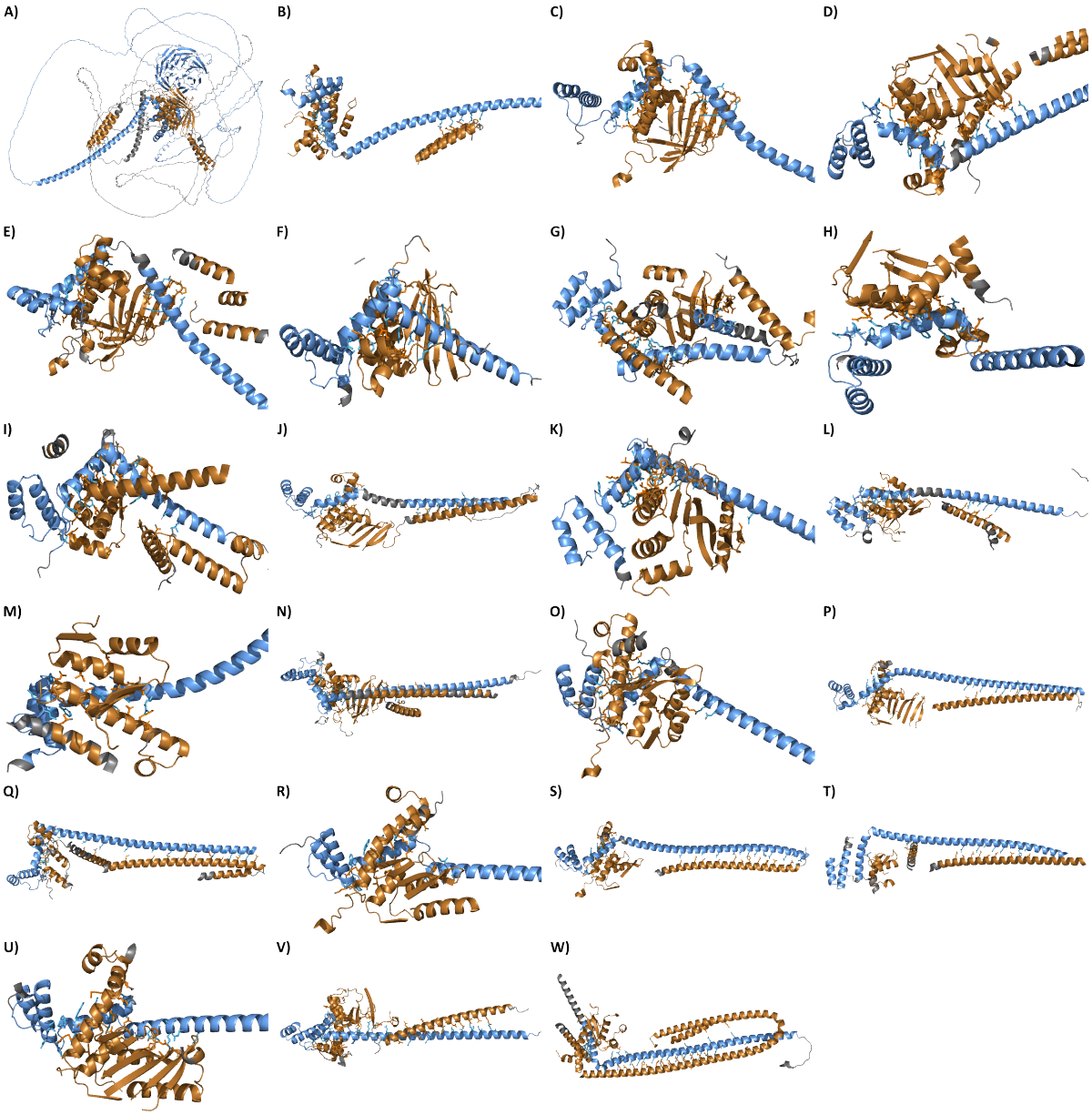
Supplementary Figure 3: Comprehensive phylogeny of the LUG family across land plants and streptophyte algae. Phylogeny with % bootstrap support per node and TOPLESS as outgroup. Phylogeny created by Clemens Rössner.



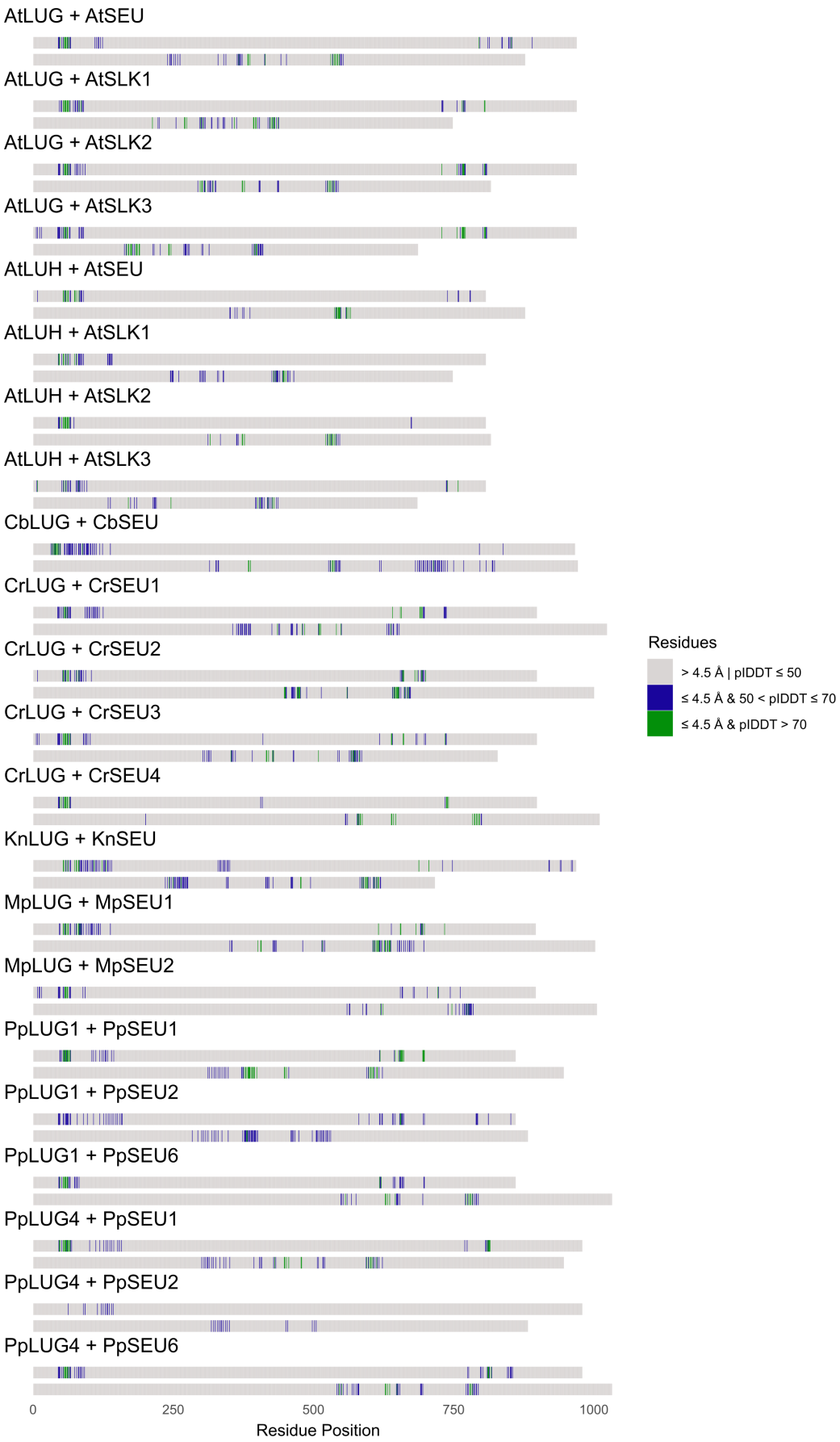
Supplementary Figure 4: Comprehensive phylogeny of the SEU family across land plants and streptophyte algae. Phylogeny with % bootstrap support per node. Phylogeny created by Clemens Rössner.

Supplementary Table 1: Mean values and standard deviation (SD) of K_a/K_s calculation. Mean and SD values for K_a and K_s scores, as well as the K_a/K_s ratios of each respective group

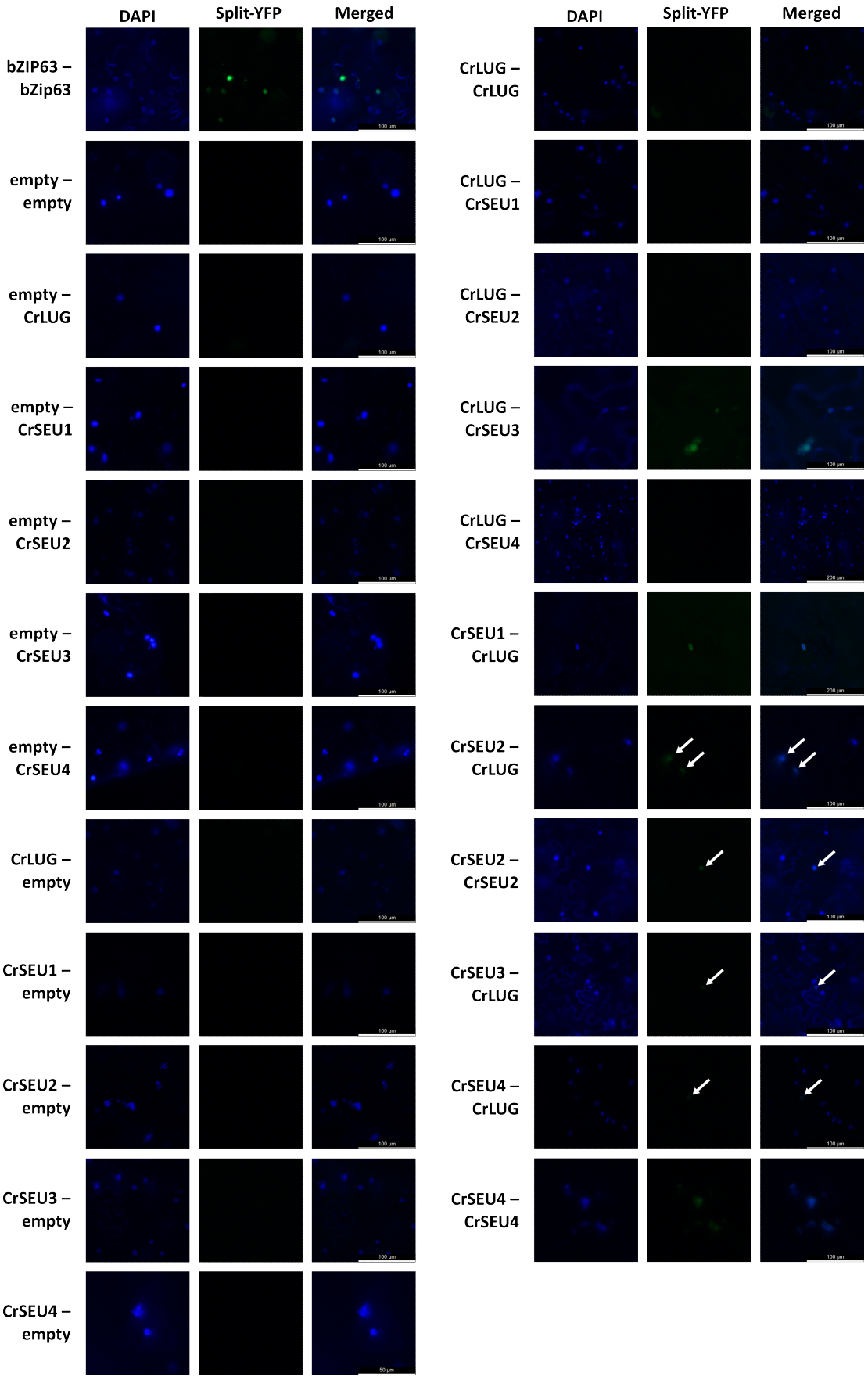
Gene	K_a		K_s		K_a/K_s	
	Mean	SD	Mean	SD	Mean	SD
LUG: Full length	0.445	0.227	2.906	0.902	0.202	0.170
LUG: LUFS	0.344	0.053	3.284	0.354	0.107	0.026
LUG: Intermediate	0.747	0.052	1.758	0.137	0.429	0.062
LUG: WD40	0.244	0.049	3.676	0.424	0.069	0.023
SEU: Full length	0.673	0.237	2.053	0.804	0.411	0.227
SEU: IDR1	0.829	0.050	1.541	0.192	0.548	0.091
SEU: LDB	0.364	0.104	3.080	0.404	0.123	0.050
SEU: IDR2	0.841	0.041	1.492	0.230	0.577	0.103
LUG – C. richardii	0.271	0.062	3.269	0.144	0.083	0.021
SEU – C. richardii	0.643	0.088	2.067	0.280	0.321	0.089



Supplementary Figure 5: Expanded *in silico* prediction of LUG – SEU dimer interaction. Overview of interacting aa residues of LUG (blue) and SEU (orange) homologs within close proximity (≤ 4.5 Å) to partner chain, with visible interacting residues. Interacting residues have a pIDDT confidence of > 50 within 4.5 Å of a residue with a pIDDT confidence of > 50 of the partner protein chain (see Parvathy *et al.*, 2024). Grey regions have a pIDDT confidence < 50. **A)** Full length AtLUG – AtSEU dimer. **B)** – **W)** First 160 N-terminal residues of LUG homologs and interacting SEU residues plus 20 surrounding atoms in both directions. **B)** AtLUG – AtSEU; **C)** AtLUG – AtSLK1; **D)** AtLUG – AtSLK2; **E)** AtLUG – AtSLK3; **F)** AtLUH – AtSEU; **G)** AtLUH – AtSLK1; **H)** AtLUH – AtSLK2; **I)** AtLUH – AtSLK3; **J)** CrLUG – CrSEU1; **K)** CrLUG – CrSEU2; **L)** CrLUG – CrSEU3; **M)** CrLUG – CrSEU4; **N)** MpLUG – MpSEU1; **O)** MpLUG – MpSEU2; **P)** PpLUG1 – PpSEU1; **Q)** PpLUG1 – PpSEU2; **R)** PpLUG1 – PpSEU6; **S)** PpLUG4 – PpSEU1; **T)** PpLUG4 – PpSEU2; **U)** PpLUG4 – PpSEU6; **V)** KnLUG – KnSEU; **W)** CbLUG – CbSEU.



Supplementary Figure 6: Expanded *in silico* mapping of interacting residues in LUG – SEU dimer. Position of interacting aa residues in LUG (upper) and SEU (lower) protein sequence. Grey residues are either not interacting, or have a pIDDT confidence of ≤ 50. Blue residues: Residues within 4.5 Å to partner chain, with pIDDT confidence between 50 and 70. Green residues: Residues within 4.5 Å to partner chain, with pIDDT confidence > 70.



Supplementary Figure 7: Extended results of independent BiFC assay of *C. richardii* CrLUG and CrSEU interactions. Positive (bZIP63) and negative (empty) controls of the BiFC assay, as well as visible YFP signals and important negative results. Results are depicted row-wise, with YC-tagged protein X – YN-tagged protein Y. All omitted interactions showed no YFP fluorescence. „DAPI“ column: Fluorescence signal of nuclei stained with 100 $\frac{\mu g}{ml}$ DAPI. „Split-YFP“ column: Split-YFP protein fluorescence. „Merged“: Overlay between DAPI and split-YFP fluorescences, with row-wise scale bars. White arrows depict difficult to see YFP fluorescence signals.

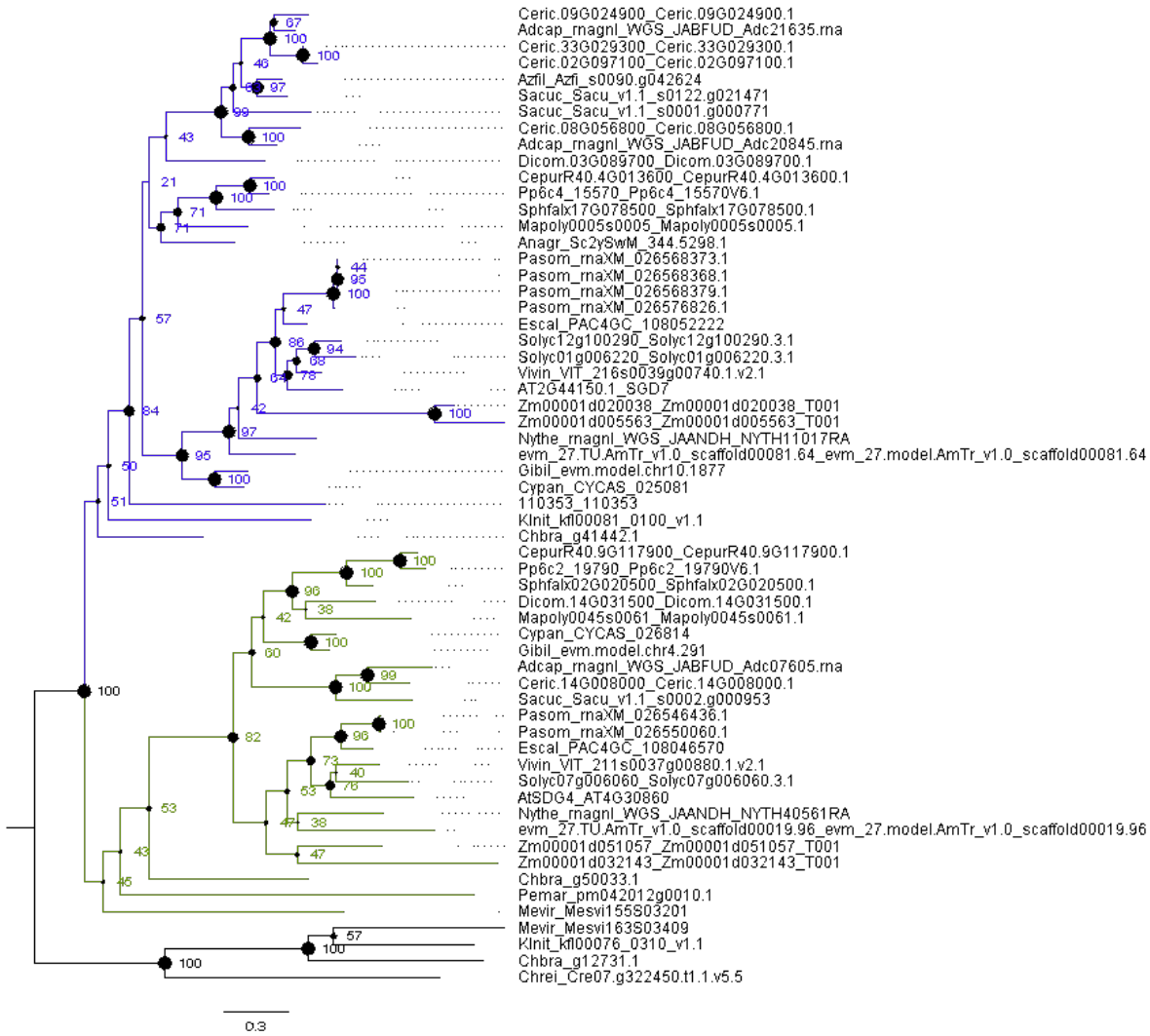
Supplementary Table 2: *t*-test of WT chloronema length and width against *pplug3 pplug4* mutant line chloronema under the influence of 5 µM exogenous NAA. Measurements of length and width taken from Fig. 19C and D, respectively. Significant differences between auxin-treated WT and mutant line chloronema were calculated using Student's *t*-test, and corrected using Benjamini–Hochberg multiple testing correction. Significance levels are displayed as asterisks or ns (Benjamini & Hochberg, 1995; Student, 1908).

WT against	Chloronema length		Chloronema width	
	t-test p-value	Signif.	*t*-test p-value	Signif.
#070	5.5×10^{-3}	**	0.33	ns
#105	3.7×10^{-7}	****	0.045	*
#106	3.2×10^{-3}	**	0.18	ns
#107	1.8×10^{-3}	**	0.19	ns
#110	1.1×10^{-4}	***	5.6×10^{-5}	****
#211	0.4	ns	0.35	ns
#212	0.97	ns	0.35	ns
#214	1.1×10^{-4}	***	6.7×10^{-4}	***

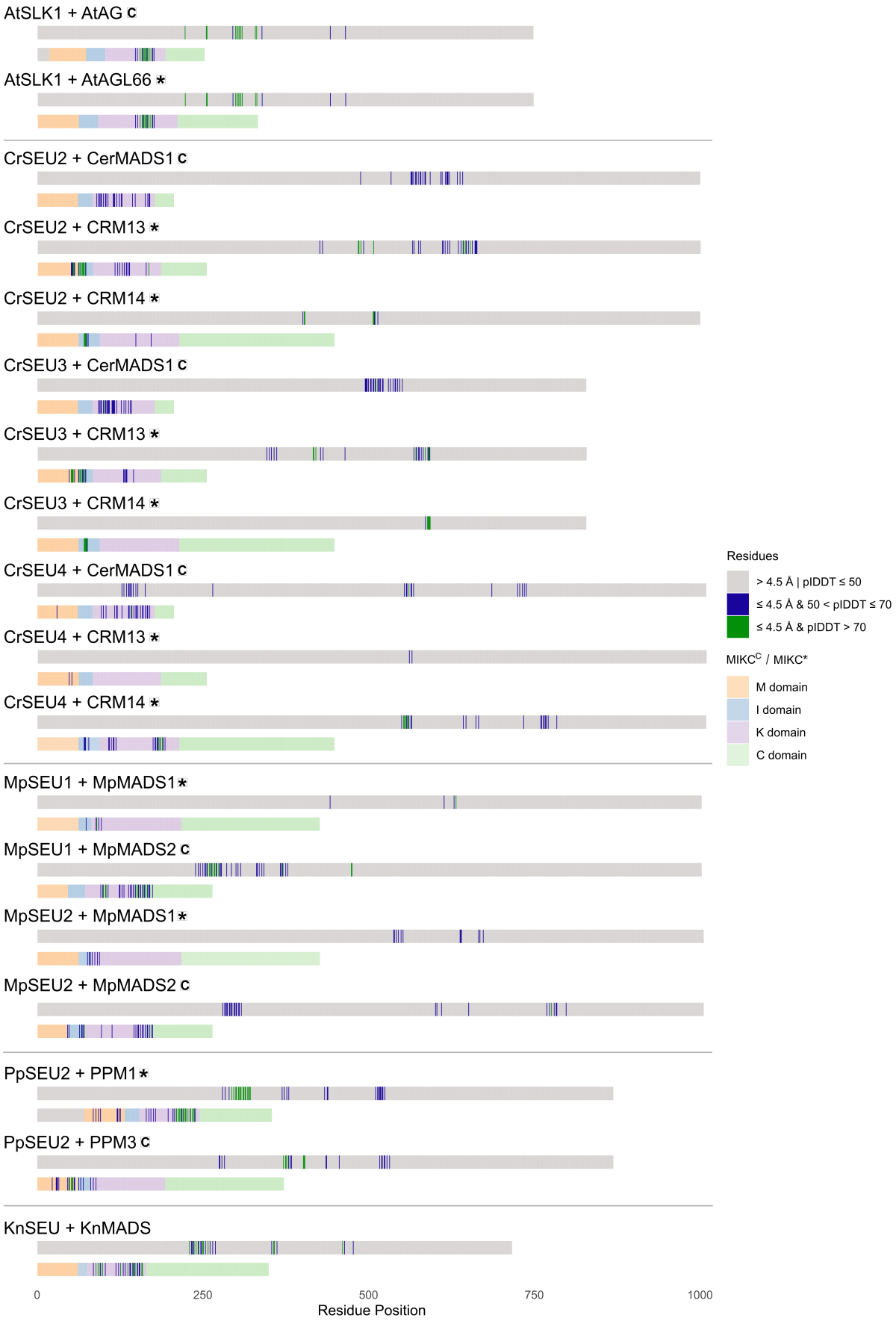
Supplementary Table 3: Selected published interactions between the LUG and SEU family members and other proteins in *A. thaliana*.

Protein	Partner	Source
LUG	CDK8	Gonzalez *et al.* (2007)
	ECAP	Shi *et al.* (2024)
	FIL	Stahle *et al.* (2009)
	HDA19	Gonzalez *et al.* (2007)
	IND	Herrera-Ubaldo *et al.* (2023)
	INO	M. K. Simon *et al.* (2017)
	JAG	Herrera-Ubaldo *et al.* (2023)
	MED14	Gonzalez *et al.* (2007)
	MED25	You *et al.* (2019)
	PHV	Herrera-Ubaldo *et al.* (2023)
	SEU	Sitaraman *et al.* (2008)
	SLK1	Stahle *et al.* (2009)
	SLK2	Stahle *et al.* (2009)
	SLK3	Stahle *et al.* (2009)
	YAB1	Herrera-Ubaldo *et al.* (2023)
	YAB2	Stahle *et al.* (2009)
	YAB5	Stahle *et al.* (2009)
LUH	FAMA	Mair *et al.* (2019)
	FIL	Stahle *et al.* (2009)
	H2B	Shrestha *et al.* (2014)
	H3	Shrestha *et al.* (2014)
	HAC1	You *et al.* (2019)
	MED25	You *et al.* (2019)
	PIF1	J. E. Lee *et al.* (2014)
	SEU	Sitaraman *et al.* (2008)
	SLK1	Shrestha *et al.* (2014)
	SLK2	Shrestha *et al.* (2014)
	SLK3	Sitaraman *et al.* (2008)
	YAB3	Stahle *et al.* (2009)
	YAB5	Stahle *et al.* (2009)

Protein	Partner	Source
SEU	AP1	Sridhar *et al.* (2004)
	AP2	Franks *et al.* (2002)
	ARR14	Herrera-Ubaldo *et al.* (2023)
	BP	Herrera-Ubaldo *et al.* (2023)
	CDK8	Gonzalez *et al.* (2007)
	ETT	Pfluger & Zambryski (2004)
	FAMA	Mair *et al.* (2019)
	GOA	Herrera-Ubaldo *et al.* (2023)
	INO	M. K. Simon *et al.* (2017)
	KNAT6	Herrera-Ubaldo *et al.* (2023)
	MED14	Gonzalez *et al.* (2007)
	PIF4	(Huai *et al.*, 2018)
	SCR	Herrera-Ubaldo *et al.* (2023)
	SDG25	Zhai *et al.* (2020)
	SDG4	Zhai *et al.* (2020)
	SEP3	Sridhar *et al.* (2004)
	SHP2	Balanzà *et al.* (2016)
	STK	Balanzà *et al.* (2016)
	STM	Herrera-Ubaldo *et al.* (2023)
	WUS	Herrera-Ubaldo *et al.* (2023)
	YAB1	Herrera-Ubaldo *et al.* (2023)
	YAB3	Stahle *et al.* (2009)
	YAB5	Stahle *et al.* (2009)
SLK2	MYB51	Frerigmann *et al.* (2014)
SLK1/2/3	YAB2	Stahle *et al.* (2009)
	YAB3	Stahle *et al.* (2009)
	YAB5	Stahle *et al.* (2009)



Supplementary Figure 8: Trimmed phylogeny of the SDG family across land plants and streptophyte algae. Phylogeny with % bootstrap support per node, with focus on the SDG4 (lime) and SDG7 (blue) homologs. Phylogeny created by Clemens Rössner.



Supplementary Figure 9: *In silico* mapping of interacting residues in SEU – MIKC dimers. Position of interacting aa residues in SEU (upper) and MADS-box (lower) protein sequence of both MIKC^C and MIKC* types, with domains marked according to Kwantes *et al.* (2012), The UniProt Consortium *et al.* (2025), & Zobell *et al.* (2010). Grey residues of SEU sequences are either not interacting, or have a pIDDT confidence of ≤ 50. Blue residues: Residues within 4.5 Å to partner chain, with pIDDT confidence between 50 and 70. Green residues: Residues within 4.5 Å to partner chain, with pIDDT confidence > 70.

Supplementary Methods

Sequence-IDs of major sequences

Supplementary Table 4: Gene-IDs and corresponding databases for principal genes used

Gene	Group	Gene-ID	Database	Genome
AtAG	MIKC^C	AT4G18960	Phytozome	*Arabidopsis thaliana* Araport11
AtAGL66	MIKC*	AT1G77980	Phytozome	*Arabidopsis thaliana* Araport11
AtAGL104	MIKC*	AT1G22130	Phytozome	*Arabidopsis thaliana* Araport11
AtLUG	LUG	AT4G32551	Phytozome	*Arabidopsis thaliana* Araport11
AtLUH	LUH	AT2G32700	Phytozome	*Arabidopsis thaliana* Araport11
AtSEP2	MIKC^C	AT3G02310	Phytozome	*Arabidopsis thaliana* Araport11
AtSEP3	MIKC^C	AT1G24260	Phytozome	*Arabidopsis thaliana* Araport11
AtSEU	SEU	AT1G43850	Phytozome	*Arabidopsis thaliana* Araport11
AtSLK1	SLK	AT4G25520	Phytozome	*Arabidopsis thaliana* Araport11
AtSLK2	SLK	AT5G62090	Phytozome	*Arabidopsis thaliana* Araport11
AtSLK3	SLK	AT4G25515	Phytozome	*Arabidopsis thaliana* Araport11
AtSTK	MIKC^C	AT4G09960	Phytozome	*Arabidopsis thaliana* Araport11
CbLUG	Algae LUG	g27887	ICIPS Garden Blaster	*Chara braunii* Cbr_1.0
CbSEU	Algae SEU	g19452	ICIPS Garden Blaster	*Chara braunii* Cbr_1.0
CerMADS1	MIKC^C	Ceric.30G027800	Phytozome	*Ceratopteris richardii* v2.1
CMADS1	MIKC^C	Ceric.29G021900	Phytozome	*Ceratopteris richardii* v2.1
CRM1	MIKC^C	Ceric.12G052800	Phytozome	*Ceratopteris richardii* v2.1
CRM13	MIKC*	Ceric.14G063000	Phytozome	*Ceratopteris richardii* v2.1
CRM14	MIKC*	Ceric.37G052700	Phytozome	*Ceratopteris richardii* v2.1
CRM16	MIKC*	Ceric.39G054900	Phytozome	*Ceratopteris richardii* v2.1
CrSEU1	SEU polyt.	Ceric.01G063700	Phytozome	*Ceratopteris richardii* v2.1
CrSEU2	SEU polyt.	Ceric.03G041400	Phytozome	*Ceratopteris richardii* v2.1
CrSEU3	SEU polyt.	Ceric.33G034400	Phytozome	*Ceratopteris richardii* v2.1

Supplementary Table 4 (cont.)

Gene	Group	Gene-ID	Database	Genome
CrSEU4	SEU/SLK-sister	Ceric.33G002300	Phytozome	*Ceratopteris richardii* v2.1
CrLUG	LUG polyt.	Ceric.35G054400	Phytozome	*Ceratopteris richardii* v2.1
KnLUG	Algae LUG	kfl01154_0010	ICIPS Garden Blaster	*Klebsormidium nitens* NIES-2145 v1.1
KnMADS	MIKC	kfl00560_0030	ICIPS Garden Blaster	*Klebsormidium nitens* NIES-2145 v1.1
KnSEU	Algae SEU	kfl00116_0070	ICIPS Garden Blaster	*Klebsormidium nitens* NIES-2145 v1.1
MpLUG	LUG polyt.	Mp1g15050	MarpolBase	MpTak-1_v7.1
MpMADS1	MIKC*	Mapoly0174s0011	Phytozome	*Marchantia polymorpha* v3.1
MpMADS2	MIKC^{C}	Mapoly0011s0161	Phytozome	*Marchantia polymorpha* v3.1
MpSEU1	SEU polyt.	Mp8g17020	MarpolBase	MpTak-1_v7.1
MpSEU2	SEU/SLK-sister	Mp1g03270	MarpolBase	MpTak-1_v7.1
PpLUG1	LUG polyt.	Pp3c1_3700	Phytozome	*Physcomitrium patens* v3.3
PpLUG2	LUG polyt.	Pp3c1_3750	Phytozome	*Physcomitrium patens* v3.3
PpLUG3	LUG polyt.	Pp6c2_17640	Phytozome	*Physcomitrium patens* v6.1
PpLUG4	LUG polyt.	Pp3c2_31650	Phytozome	*Physcomitrium patens* v3.3
PPM1	MIKC^{C}	Pp3c12_22500	Phytozome	*Physcomitrium patens* v3.3
PPM2	MIKC^{C}	Pp3c3_31980	Phytozome	*Physcomitrium patens* v3.3
PPM3	MIKC*	Pp3c14_22180	Phytozome	*Physcomitrium patens* v3.3
PPM9	MIKC*	Pp3c8_6920	Phytozome	*Physcomitrium patens* v3.3
PPMADS1	MIKC^{C}	Pp3c4_540	Phytozome	*Physcomitrium patens* v3.3
PPMADS2	MIKC*	Pp3c12_22890	Phytozome	*Physcomitrium patens* v3.3
PpSEU1	SEU polyt.	Pp3c25_7810	Phytozome	*Physcomitrium patens* v3.3
PpSEU2	SEU polyt.	Pp3c25_7960	Phytozome	*Physcomitrium patens* v3.3
PPTIM2	M-type MADS	Pp3c16_19170	Phytozome	*Physcomitrium patens* v3.3
PPTIM3	M-type MADS	Pp3c25_6940	Phytozome	*Physcomitrium patens* v3.3

Note: This table only contains IDs for those sequences used in the Y2H and BiFC assays, or those used for closer *in silico* analyses like interacting protein residues and transcriptome analysis. See section 3.2.2 for information about additional sequences.

References:

- ICIPS Garden Blaster: Priyam *et al.* (2019) & Roessner *et al.* (2024)
- MarpolBase: Tanizawa *et al.* (2025)
- Phytozome: Goodstein *et al.* (2012)
- *Arabidopsis thaliana* Araport11: Cheng *et al.* (2017)
- *Ceratopteris richardii* v2.1: D. B. Marchant *et al.* (2022)
- *Chara braunii* Cbr_1.0: Nishiyama *et al.* (2018)
- *Klebsormidium nitens* NIES-2145 v1.1: Hori *et al.* (2014)
- *Marchantia polymorpha* v3.1: Bowman *et al.* (2017)
- MpTak-1_v7.1: Tanizawa *et al.* (2025)
- *Physcomitrium patens* v3.3: Lang *et al.* (2018)
- *Physcomitrium patens* v6.1: Bi *et al.* (2024)

Plasmids

Supplementary Table 5: Used plasmids containing *A. thaliana* sequences

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pDEST22	*AtAGL66* CDS	1454	*E.coli* Dh5α	Ampicillin	Folter *et al.* (2005)
pDEST22	*AtAGL104* CDS	1455	*E.coli* Dh5α	Ampicillin	Folter *et al.* (2005)
pDONR221	*AtAG* CDS	1448	*E.coli* Dh5α	Kanamycin	Folter *et al.* (2005)
pGADT7	*AtAG* CDS	1494	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*AtAGL66* CDS	1480	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*AtAGL104* CDS	1472	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*AtLUG* CDS	1469	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*AtLUH* CDS	1470	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*AtSEP2* CDS	1403	*E.coli* Dh5α	Ampicillin	Gift by J. Golz
pGADT7	*AtSEP3* CDS	1407	*E.coli* Dh5α	Ampicillin	Gift by J. Golz
pGADT7	*AtSEU* CDS	1399	*E.coli* Dh5α	Ampicillin	Stahle *et al.* (2009)
pGADT7	*AtSLK1* CDS	1464	*E.coli* Dh5α	Ampicillin	Stahle *et al.* (2009)
pGADT7	*AtSLK2* CDS	1401	*E.coli* Dh5α	Ampicillin	Stahle *et al.* (2009)
pGADT7	*AtSLK3* CDS	1402	*E.coli* Dh5α	Ampicillin	Stahle *et al.* (2009)
pGADT7	*AtSTK* CDS	1406	*E.coli* Dh5α	Ampicillin	Gift by J. Golz
pGBKT7	*AtAG* CDS	1495	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*AtAGL66* CDS	1481	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*AtAGL104* CDS	1473	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*AtLUG* CDS	1419	*E.coli* Dh5α	Kanamycin	Stahle *et al.* (2009)
pGBKT7	*AtLUH* CDS	1420	*E.coli* Dh5α	Kanamycin	Stahle *et al.* (2009)
pGBKT7	*AtSEP2* CDS	1474	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*AtSEP3* CDS	1426	*E.coli* Dh5α	Kanamycin	Gift by J. Golz
pGBKT7	*AtSEU* CDS	1421	*E.coli* Dh5α	Kanamycin	Stahle *et al.* (2009)
pGBKT7	*AtSLK1* CDS	1422	*E.coli* Dh5α	Kanamycin	Stahle *et al.* (2009)
pGBKT7	*AtSLK2* CDS	1423	*E.coli* Dh5α	Kanamycin	Stahle *et al.* (2009)

Supplementary Table 5 (cont.)

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pGBKT7	*AtSLK3* CDS	1471	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*AtSTK* CDS	1475	*E. coli* JM109	Kanamycin	This dissertation

Supplementary Table 6: General plasmids used

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
p19	-	280	*A. tumefaciens* GV3101	Kanamycin Gentamycin	-
pENTR4	-	815	*E. coli* DB3.1	Kanamycin	Thermo Fisher
pGADT7	-	51	*E.coli* Dh5α	Ampicillin	Clontech (2008)
pGADT7	*EScaGLO* CDS	160	*E. coli* DH10B	Ampicillin	Lange *et al.* (2013)
pGADT7-GW	-	1110	*E. coli* DB3.1	Ampicillin	Lu *et al.* (2010)
pGBKT7	-	52	*E.coli* Dh5α	Kanamycin	Clontech (2008)
pGBKT7	*EcDEF1* CDS	163	*E. coli* DH10B	Kanamycin	Lange *et al.* (2013)
pGBKT7-GW	-	1111	*E. coli* DB3.1	Kanamycin	Lu *et al.* (2010)
pGTQL1211YN	-	1112	*E. coli* DB3.1	Kanamycin	Lu *et al.* (2010)
pGTQL1221YC	-	1113	*E. coli* DB3.1	Kanamycin	Lu *et al.* (2010)
pMLBART	35S::bZIP63-YC	413	*A. tumefaciens* GV3101	Spectino-mycin Gentamycin	Lange *et al.* (2013)
pMLBART	35S::bZIP63-YN	414	*A. tumefaciens* GV3101	Spectino-mycin Gentamycin	Lange *et al.* (2013)
pMK-Cas9-gate	-		*E. coli* DB3.1	Ampicillin Neomycin	Mallett *et al.* (2019)
pTNT	-	54	*E.coli* Dh5α	Ampicillin	Promega, Madison, USA
pENTR-PpU6P-sgRNA-L1L4	-	1171	*E.coli* JM109	Kanamycin	Mallett *et al.* (2019)

Supplementary Table 6 (cont.)

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pENTR-PpU6P-sgRNA-L3L2	-	1175	*E.coli* JM109	Kanamycin	Mallett *et al.* (2019)
pENTR-PpU6P-sgRNA-R4R3	-	1174	*E.coli* JM109	Kanamycin	Mallett *et al.* (2019)

Supplementary Table 7: Used plasmids containing *C. richardii* sequences

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pENTR4	*CerMADS1* CDS	1393	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*CMADS1* CDS	1394	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*CRM1* CDS	1392	*E.coli* JM109	Kanamycin	This dissertation
pGADT7	*CRM13* (codon optimized CDS)	1395	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)
pGADT7	*CRM14* (codon optimized CDS)	1396	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)
pGADT7	*CRM16* (codon optimized CDS)	1397	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)
pGADT7-GW	*CerMADS1* CDS	1398	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CMADS1* CDS	1439	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CrLUG* CDS	1193	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CRM1* CDS	1438	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CrSEU1* CDS	1194	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CrSEU2* CDS	1195	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CrSEU3* CDS	1196	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*CrSEU4* CDS	1197	*E.coli* Dh5α	Ampicillin	This dissertation
pGBKT7	*CRM1* CDS	1477	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7	*CRM13* (codon optimized CDS)	1388	*E.coli* Dh5α	Kanamycin	Rümpler *et al.* (2023)

Supplementary Table 7 (cont.)

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pGBKT7	*CRM14* (codon optimized CDS)	1390	*E.coli* Dh5α	Kanamycin	Rümpler *et al.* (2023)
pGBKT7	*CRM16* (codon optimized CDS)	1391	*E.coli* Dh5α	Kanamycin	Rümpler *et al.* (2023)
pGBKT7-GW	*CerMADS1* CDS	1482	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CMADS1* CDS	1446	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CrLUG* CDS	1199	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CrSEU1* CDS	1200	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CrSEU2* CDS	1201	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CrSEU3* CDS	1227	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*CrSEU4* CDS	1228	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1211YN	*CrLUG* CDS	1231	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1211YN	*CrSEU1* CDS	1232	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1211YN	*CrSEU2* CDS	1233	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1211YN	*CrSEU3* CDS	1234	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1211YN	*CrSEU4* CDS	1235	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1221YC	*CrLUG* CDS	1190	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1221YC	*CrSEU1* CDS	1191	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1221YC	*CrSEU2* CDS	1202	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1221YC	*CrSEU3* CDS	1198	*E.coli* Dh5α	Kanamycin	This dissertation
pGTQL1221YC	*CrSEU4* CDS	1192	*E.coli* Dh5α	Kanamycin	This dissertation
pTNT	*CRM13* (codon optimized CDS)	1371	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)
pTNT	*CRM14* (codon optimized CDS)	1372	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)
pTNT	*CRM16* (codon optimized CDS)	1374	*E.coli* Dh5α	Ampicillin	Rümpler *et al.* (2023)

Supplementary Table 8: Used plasmids containing *K. nitens* sequences

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pGADT7	*KnLUG* CDS	1562	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*KnMADS* CDS	1564	*E. coli* JM109	Ampicillin	This dissertation
pGADT7	*KnSEU* CDS	1563	*E. coli* JM109	Ampicillin	This dissertation
pGBKT7	*KnLUG* CDS	1560	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*KnMADS* CDS	1559	*E. coli* JM109	Kanamycin	This dissertation
pGBKT7	*KnSEU* CDS	1561	*E. coli* JM109	Kanamycin	This dissertation
pMA	*KnLUG* CDS	1553	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMA	*KnMADS* CDS	1554	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMA	*KnSEU* CDS	1552	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis

Supplementary Table 9: Used plasmids containing *M. polymorpha* sequences

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pGADT7	*MpLUG* CDS	1414	*E.coli* Dh5α	Ampicillin	Q. Li (2022)
pGADT7	*MpMADS1* CDS	1417	*E.coli* Dh5α	Ampicillin	Q. Li (2022)
pGADT7	*MpMADS2* CDS	1418	*E.coli* Dh5α	Ampicillin	Q. Li (2022)
pGADT7	*MpSEU1* CDS	1486	*E.coli* JM109	Ampicillin	Q. Li (2022)
pGADT7	*MpSEU2* CDS	1416	*E.coli* Dh5α	Ampicillin	Q. Li (2022)
pGBKT7	*MpLUG* CDS	1428	*E.coli* Dh5α	Kanamycin	Q. Li (2022)
pGBKT7	*MpMADS1* CDS	1431	*E.coli* Dh5α	Kanamycin	Q. Li (2022)
pGBKT7	*MpMADS2* CDS	1432	*E.coli* Dh5α	Kanamycin	Q. Li (2022)
pGBKT7	*MpSEU1* CDS	1429	*E.coli* Dh5α	Kanamycin	Q. Li (2022)
pGBKT7	*MpSEU2* CDS	1430	*E.coli* Dh5α	Kanamycin	Q. Li (2022)

Supplementary Table 10: Used plasmids containing *P. patens* sequences

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pENTR4	*PpLUG1* CDS	1253	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*PpLUG2*_synthesis	1328	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PpLUG3*_synthesis	1329	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PpLUG4* CDS	1243	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*PPM1* CDS	1357	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PPM2* CDS	1358	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*PPM3* CDS	1353	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PPM9* CDS	1354	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PPMADS1* CDS	1355	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*PPMADS2* CDS	1361	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PpSEU*_synthesis	1350	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PpSEU2*_synthesis	1335	*E.coli* Dh5α	Kanamycin	This dissertation
pENTR4	*PPTIM2* CDS	1359	*E.coli* JM109	Kanamycin	This dissertation
pENTR4	*PPTIM3* CDS	1360	*E.coli* JM109	Kanamycin	This dissertation
pGADT7-GW	*PpLUG1* CDS	1255	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PpLUG2*_synthesis	1339	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PpLUG3*_synthesis	1337	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PpLUG4* CDS	1256	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPM1* CDS	1362	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPM2* CDS	1363	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPM3* CDS	1365	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPM9* CDS	1366	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPMADS1* CDS	1364	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPMADS2* CDS	1376	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PpSEU*_synthesis	1351	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PpSEU2*_synthesis	1338	*E.coli* Dh5α	Ampicillin	This dissertation

Supplementary Table 10 (cont.)

Vector	Insert	Stock-No.	Bacterial strain	Resistance	Source
pGADT7-GW	*PPTIM2* CDS	1377	*E.coli* Dh5α	Ampicillin	This dissertation
pGADT7-GW	*PPTIM3* CDS	1367	*E.coli* Dh5α	Ampicillin	This dissertation
pGBKT7-GW	*PpLUG1* CDS	1265	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PpLUG2*_synthesis	1340	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PpLUG3*_synthesis	1341	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PpLUG4* CDS	1264	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPM1* CDS	1379	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPM2* CDS	1368	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPM3* CDS	1380	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPM9* CDS	1381	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPMADS1* CDS	1382	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPMADS2* CDS	1378	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PpSEU*_synthesis	1352	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PpSEU2*_synthesis	1342	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPTIM2* CDS	1369	*E.coli* Dh5α	Kanamycin	This dissertation
pGBKT7-GW	*PPTIM3* CDS	1381	*E.coli* Dh5α	Kanamycin	This dissertation
pMA	*PpLUG2*_synthesis	1280	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMA	*PpLUG3*_synthesis	1281	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMA	*PpSEU*_synthesis	1326	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMA	*PpSEU2*_synthesis	1327	*E.coli* Dh5α	Ampicillin	GeneArt Custom Gene Synthesis
pMK-Cas9-gate	*PpLUG3/4* Protospacer	1204	*E. coli* Dh5α	Ampicillin Neomycin	This dissertation

Primer sequences

Supplementary Table 11: Cloning primers for Y2H constructs

CDS	Name	Adapter + RE site	Complementary sequence
AtAG	AtAG_XmaI_fw_new	ACGTCCCGGGA	ATGGCGTACCAATCGGAG
	AtAG_BamHI_rv	ACGTGGATCC	TTACACTAACTGGAGAGCGG
AtAGL66	AtAGL66_XmaI_fw	ACGTCCCGGGA	ATGGGTCGAGTGAAATTGGA
	AtAGL66_BamHI_rv	ACGTGGATCC	TTATTGGCTACTGAGCTGAGG
	AtAGL66_NotI_rv	ACATGCGGCCGC	TTATTGGCTACTGAGCTGAGG
AtAGL104	AtAGL104_XmaI_fw	ACGTCCCGGGA	ATGGGTCGGGTGAAATTAGA
	AtAGL104_NotI_rv	ACATGCGGCCGC	TTATTGGCTACTGAGTTGAGGA
	AtAGL104_XhoI_rv	ACTGCTCGAG	TTATTGGCTACTGAGTTGAGGA
AtLUG	AtLUG_XmaI_fw	ACGTCCCGGGA	ATGTCTCAGACCAACTGGG
	AtLUG_XhoI_rv	ACTGCTCGAG	TCACTTCCACAGTTTCACTAG
AtLUH	AtLUH_XmaI_fw	ACGTCCCGGGA	ATGGCTCAGAGTAATTGGGA
	AtLUH_XhoI_rv	ACTGCTCGAG	CTACTTCCAAATCTTTACGGATT
AtSEP2	AtSEP2_XmaI_fw	ACGTCCCGGGA	ATGGGAAGAGGAAGAGTAGAG
	AtSEP2_NotI_rv	ACATGCGGCCGC	TCACAGCATCCAGCCAG
AtSLK3	AtSLK3_XmaI_fw	ACGTCCCGGGA	ATGCAGAGGAGCAGTGGC
	AtSLK3_NotI_rv	ACATGCGGCCGC	TTACAAGCCACCATAGATATCACTG
AtSTK	AtSTK_XmaI_fw	ACGTCCCGGGA	ATGCTCTTTCCCCATGAAAG
	AtSTK_NotI_rv	ACATGCGGCCGC	TTAGATAGATGTTCTCATGTCAAG
CerMADS1	CerMADS1_NotI_fw	ACATGCGGCCGC	ATGGTGAGGACGAAGATCAAGATTAAGCG
	CerMADS1_NotI_rv	ACATGCGGCCGC	TCACCGTCCCGCGCGTG
CMADS1	CMADS1_KpnI_fw	ACATGGTACCCA	ATGAAGCAGCTTCCGCACA
	CMADS1_NotI_rv	ACATGCGGCCGC	CTACAACCTCAGGTCAAGTTTC
CrLUG	Y-B_LUG_KpnI_fw	ACATGGTACCCA	ATGGCACAAACAAAATGGGAA
	Y2H_LUG_NotI_rv	ACATGCGGCCGC	CTACTTCCATAACTTGACCATCTTA

Supplementary Table 11 (cont.)

CDS	Name	Adapter + RE site	Complementary sequence
CRM1	CRM1_KpnI_fw	ACATGGTACCCA	ATGGTGAGGAGGAAGATCAA
	CRM1_XmaI_fw	ACGTCCCGGGA	ATGGTGAGGAGGAAGATCAA
	CRM1_NotI_rv	ACATGCGGCCGC	CTACTCAGAGGCAGGTGATT
CrSEU1	Y-B_SEU1_KpnI_fw	ACATGGTACCCA	ATGGATTCAAGCCATCAGGCA
	Y2H_SEU1_NotI_rv	ACATGCGGCCGC	TTATGGGGATTTCCAGCTGAATG
CrSEU2	Y-B_SEU2_KpnI_fw	ACATGGTACCCA	ATGGATTCAAGCCACCGGG
	Y2H_SEU2_NotI_rv	ACATGCGGCCGC	TTACGGAGACTTCCAGCTAAATG
CrSEU3	Y-B_SEU3_KpnI_fw	ACATGGTACCCA	ATGGCAGGTTCTGAGAGTGA
	Y2H_SEU3_NotI_rv	ACATGCGGCCGC	TTACTGCTGATGAGTCTCAGAAATG
CrSEU4	Y-B_SEU4_KpnI_fw	ACATGGTACCCA	ATGATGGACATGCGGCAGC
	Y2H_SEU4_NotI_rv	ACATGCGGCCGC	CTAGGTATTAGTTCTCATGTCACTCATCTC
PpLUG1	PpLUG_KpnI_fw	ACATGGTACCCA	ATGGCGCAGAGCAACTG
	PpLUG_NotI_rv	ACATGCGGCCGC	CTACTTCCATAACCTCACGGA
PpLUG2	PpLUG2_KpnI_fw	ACATGGTACCCA	ATGGCGCAGAGTAACTGG
	PpLUG_NotI_rv	ACATGCGGCCGC	CTACTTCCATAACCTCACGGA
PpLUG3	PpLUG3_KpnI_fw	ACATGGTACCCA	ATGGCTCAGAGCAACTGG
	PpLUG3_NotI_rv	ACATGCGGCCGC	TCATTTCCATAACCTCACCGT
PPM1	PPM1_KpnI_fw	ACATGGTACCCA	ATGATGACTGCGGCGTGTGGA
	PPM1_NotI_rv	ACATGCGGCCGC	TTAAATTCTCACACTTGGTAAAATG
PPM2	PPM2_KpnI_fw	ACATGGTACCCA	ATGACTGCGTCTGGACAATC
	PPM2_NotI_rv	ACATGCGGCCGC	CTAAATTCTCACACTTGGTGA
PPM3	PPM3_KpnI_fw	ACATGGTACCCA	ATGGGCAGGGTAAAGCTAG
	PPM3_NotI_rv	ACATGCGGCCGC	TCATTTCCAAGCCTCTTCATT
PPM9	PPM9_KpnI_fw	ACATGGTACCCA	ATGGGGAGGGTTAAATTGGA
	PPM9_NotI_rv	ACATGCGGCCGC	TCATTTCCACGCATGATGTC
PPMADS1	PPMADS1_KpnI_fw	ACATGGTACCCA	ATGATGACTACAGCGTCTG
	PPMADS1_NotI_rv	ACATGCGGCCGC	TTATCGATCGACATGTAACTTA

Supplementary Table 11 (cont.)

CDS	Name	Adapter + RE site	Complementary sequence
PPMADS2	PPMADS2_SalI_fw	ACATGTCGAC	ATGGGGAGGGTTAAGCTG
	PPMADS2_SalI_rv	ACATGTCGAC	TTACTTCCACGCAGGATGG
PpSEU1	PpSEU_NotI_fw	ACATGCGGCCGC	ATGGCTGGTGCTCACGC
	PpSEU_NotI_rv	ACATGCGGCCGC	CTAGTTGAACAGCCCTCCAAAC
PpSEU2	PpSEU2_KpnI_fw	ACATGGTACCCA	ATGGAGAGGGGCGATGG
	PpSEU_NotI_rv	ACATGCGGCCGC	CTAGTTGAACAGCCCTCCAAAC
PPTIM2	PPTIM2_KpnI_fw	ACATGGTACCCA	ATGGGGCGTGCTAAGATAGAG
	PPTIM2_NotI_rv	ACATGCGGCCGC	TCACGACAAGAGCTTTTCTCC
PPTIM3	PPTIM3_KpnI_fw	ACATGGTACCCA	ATGGGTCGTGCCAAGATTGA
	PPTIM3_NotI_rv	ACATGCGGCCGC	TTACCCAGGCAACTCTAGTTTC

Supplementary Table 12: Cloning primers for BiFC constructs

CDS	Name	Adapter + RE site	Complementary sequence
CrLUG	Y-B_LUG_KpnI_fw	ACATGGTACCCA	ATGGCACAAACAAAATGGGAA
	BFC_LUG_NotI_rv	ACATGCGGCCGCCA	CTTCCATAACTTGACCATCTTATC
CrSEU1	Y-B_SEU1_KpnI_fw	ACATGGTACCCA	ATGGATTCAAGCCATCAGGCA
	BFC_SEU1_NotI_rv	ACATGCGGCCGCCA	TGGGGATTTCCAGCTGAATGGA
CrSEU2	Y-B_SEU2_KpnI_fw	ACATGGTACCCA	ATGGATTCAAGCCACCGGG
	BFC_SEU2_NotI_rv	ACATGCGGCCGCCA	CGGAGACTTCCAGCTAAATGGC
CrSEU3	Y-B_SEU3_KpnI_fw	ACATGGTACCCA	ATGGCAGGTTCTGAGAGTGA
	BFC_SEU3_NotI_rv	ACATGCGGCCGCCA	CTGCTGATGAGTCTCAGAAATGA
CrSEU4	Y-B_SEU4_KpnI_fw	ACATGGTACCCA	ATGATGGACATGCGGCAGC
	BFC_SEU4_NotI_rv	ACATGCGGCCGCCA	GGTATTAGTTCTCATGTCACTCATCT

Supplementary Table 13: Cloning primers for *P. patens* mutagenesis

CDS	Name	Adapter	Protospacer sequence
PpLUG3	PpLUG3_sgRNA1_fw	ccat	ACACTATGGCTCAGAGCAAC
	PpLUG3_sgRNA1_rv	aaac	GTTGCTCTGAGCCATAGTGT
PpLUG3 / *PpLUG4*	PpLUG3/4_sgRNA2_fw	ccat	CATTCGAACAAAAACCCTCC
	PpLUG3/4_sgRNA2_rv	aaac	GGAGGGTTTTTGTTCGAATG
PpLUG4	PpLUG4_sgRNA1_fw	ccat	AAAGTGTCGTCCGACCCCGT
	PpLUG4_sgRNA1_rv	aaac	ACGGGGTCGGACGACACTTT

Supplementary Table 14: CDS sequencing primers

Gene	Name	Sequence
CRM13	CRM13_rv	TCACAACAGATTGCTTTGCATG
CRM14	CRM14_rv	TCAGTTGTAAGACGAGCCGC
CRM16	CRM16_rv	TTAAGTATTGTCATTGAACCATGC
CrLUG	CrLUG_Seq_Int_fw	TCAGCAGGCTCAAGGAAGT
	CrLUG_Seq_rv	GGTTGCGTGCCTGTATTTGCT
CrSEU1	CrSEU1_Seq_Int_1_fw	GCCTTTGCATACCCAAATTA
	CrSEU1_Seq_Int_2_fw	ACGAAGGGTTTTAGTTCC
	CrSEU1_Seq_rv	GAAGCTGCTGAGAATCTTGTC
CrSEU2	CrSEU2_Seq_Int_1_fw	CACACTCAGATTTCTATACC
	CrSEU2_Seq_Int_2_fw	CTACACTAAAAGATATGTTCGA
	CrSEU2_Seq_rv	GCTCTAACTGTTGCTGTTGAGA
CrSEU3	CrSEU3_Seq_Int_fw	CTGCAGCAACTGCAGCAA
	CrSEU3_Seq_rv	CAGGTTTGTATCCATTCTCCAAT
	CrSEU3_Outer_fw	AAGCAAAGAAAGCAACTGCAAGT
	CrSEU3_Outer_rv	TGTTACTCCTCTGGCTACCTGA
CrSEU4	CrSEU4_Seq_Int_fw	TGGACATGGAGAGAATGTT
	CrSEU4_Seq_rv	AAAGAATGTCCTCTCGCATTGG

Supplementary Table 14 (cont.)

Gene	Name	Sequence
PpLUG	PpLUG_Seq_int_fw	ACATGTCAACGTCAGGTTCTC
	PpLUG_Seq_int2_fw	GATGTGCTGAACCAAAGGTCT
	PpLUG_PCR_rv	GTGCTTCAGCCGCTCTTCG
PpLUG2	PpLUG2_PCR_rv	AACAAGTGCATCAACAGCAAGC
PpLUG3	PpLUG3_PCR_rv	CAACAGCTTCCACAGCGCA
PpLUG4	PpLUG4_Seq1_int_fw	GCGCGCATTGTATCGGAT
	PpLUG4_Seq2_int_fw	GTTGCCCCAAGTGCAAGAAA
	PpLUG4_Seq3_int_fw	TTGCTGGGAATGCAACTACA
PpLUG4	PpLUG4_PCR_rv	TCTCCGTTGTTGTGCATGC
PpSEU	PpSEU_SEU2_Seq_int_fw	ACAGATACAGGCCCATCTGA
	PpSEU_PCR_rv	GTTCAGGTACACCTCCGCTA
PpSEU2	PpSEU_SEU2_Seq_int_fw	ACAGATACAGGCCCATCTGA
	PpSEU2_PCR_rv	TGCTGCTGCTCCATTCGC

Supplementary Table 15: Sequencing primers for plasmid backbones

Vector	Name	Sequence
pENTR4	pENTR_Seq_fw	TTCTACAAACTCTTCCTGTTAG
	pENTR_Seq_rv	TGCAATGTAACATCAGAGATT
	pENTR4-5'MCS_Seq	ATTGGTTCCCATGGTGGAG
pENTR-PpU6P-sgRNA-L1L4	M13_fw	GTAAAACGACGGCCAGT
	M13_rv	CAGGAAACAGCTATGAC
pENTR-PpU6P-sgRNA-R4R3	M13_fw	GTAAAACGACGGCCAGT
	M13_rv	CAGGAAACAGCTATGAC
pENTR-PpU6P-sgRNA-L3L2	M13_fw	GTAAAACGACGGCCAGT
	M13_rv	CAGGAAACAGCTATGAC
pGADT7-GW	AD_rev	CGATGCACAGTTGAAGTGAACT
	pGADT7-GW_Seq_fw	AAGCGGAATTAATTCCCGAG

Supplementary Table 15 (cont.)

Vector	Name	Sequence
pGBKT7-GW	BD_rev	CCGGAATTAGCTTGGCTGC
	pGBKT7-GW_Seq_fw	GACTGGAACAGCTATTTCTACTGAT
pGTQL1211YN	pGTQL1211YN_Seq_rv	CTCTAGACTCACCTAGATTAACTCT
pGTQL1221YC	pGTQL1221YC_Seq_rv	CTCTCTAGACTCACCTAGAATTAA
pMK-Cas9-gate	pUbi_rv	GACCGGGTCACGCTGCACTG
pTNT	Sp6_pTNT1_fw	ATTTAGGTGACACTATAGAATAC
	Sp6_pTNT2_fw	TAGCATTTAGGTGACACTATAG
	pTNT_rv2	TTTTTTTTTGCGGCCGCCC
	pTNT_rv3	CTTCGCTATTACGCCAGC

Supplementary Table 16: Genotyping primers for *P. patens* mutagenesis

Gene	Name	Sequence
PpLUG3	PpLUG3_sgRNA1_Seq_fw	TCGAATGGAACTTGCCGA
	PpLUG3_sgRNA1_Seq_rv	CTCATTCATTTTAGTACAGTCTG
	PpLUG3_sgRNA1_New_fw	AGTGCGAAGCAAACCCTGG
	PpLUG3_sgRNA1_New_rv	ATTCGCCTCAGCTCTTCTAGC
	PpLUG3_sgRNA2_New_fw	GTTGTGCCGGGATTGCATG
	PpLUG3_sgRNA2_New_rv	TACGCCTCTACCATGCCACC
	PpLUG3_Off_fw	GGTTTGATGTGCAGGCTGGA
PpLUG3 / *PpLUG4*	PpLUG3/4_sgRNA2_Seq_fw	TTGTGGGAAGGTAGTGGAC
	PpLUG3_sgRNA2_Seq_rv	AACCACAACACCCACCCTA
	PpLUG4_sgRNA2_Seq_rv	CTCGTGCTCCCTACCTCA

Supplementary Table 16 (cont.)

Gene	Name	Sequence
PpLUG4	PpLUG4_sgRNA1_Seq_fw	AAGTTGGTTGGGGAGGGT
	PpLUG4_sgRNA1_Seq_rv	CGCATCACAGCAACACCA
	PpLUG4_sgRNA1_New_fw	AATGGGCTCGGTAGTGTGTGA
	PpLUG4_sgRNA1_New2_fw	CTGGAAAATTCTGAGCCGCTCGA
	PpLUG4_sgRNA1_New_rv	CCCATCCACCCTTTCCCTCA
	PpLUG4_sgRNA2_New_fw	ATGGGTGGGTGTCGTTGTGC
	PpLUG4_sgRNA2_New_rv	CCCTTCCTGCACTCTCCAAC
	PpLUG4_Outer_fw	GAAGGCTTGGTGGTCGGT
	PpLUG4_Off_rv	ACCCTCCCAGTCCTTACGAC

All primers used were synthesized by Eurofins Genomics.

Supplementary Scripts

Scripts written and used for automated analysis steps in this thesis.

Note: For the sake of brevity and clarity, R scripts used purely for statistical analysis and simple data visualization in ggplot2 were not included here. This mainly concerns sections 3.2.6, 3.2.4, and 3.5.4

All supplementary scripts can also be found in the following GitHub repository:

https://github.com/Garrecht/800-Million-years-of-co-evolution-in-the-green-plant-lineage-the-case-of-LEUNIG-and-SEUSS.git

```python
import os
import csv
import pymol
from pymol import cmd

def collapse_residue_ranges(residues):
    """Collapse a list of domain positions into residue ranges."""
    if not residues:
        return []

    residues = sorted(map(int, residues))  # Convert residues to integers
    collapsed = []

    start = residues[0]
    end = residues[0]

    for i in range(1, len(residues)):
        if residues[i] == end + 1:
            end = residues[i]
        else:
            if start == end:
                collapsed.append((start, end))  # Single residue
            else:
                collapsed.append((start, end))  # Range
            start = residues[i]
            end = residues[i]

    # Append the last range or single residue
    collapsed.append((start, end))
    return collapsed

def extract_helix_positions(cif_file, threshold=70):
    """Extract alpha-helix positions for chains with pIDDT >= threshold."""
    cmd.load(cif_file)

    helix_data = []
    for chain in cmd.get_chains():
        cmd.select(f'chain_{chain}', f'chain {chain}')

        atoms = cmd.get_model(f'chain_{chain}')
        selected_residues = [atom.resi for atom in atoms.atom if atom.b >= threshold]

        if selected_residues:
            residue_selection = '+'.join(map(str, selected_residues))
            cmd.select(f'helix_{chain}', f'chain {chain} and ss H and resi {
            residue_selection}')

            helix_residues = cmd.get_model(f'helix_{chain}')
            if helix_residues:
                helix_residue_numbers = [int(atom.resi) for atom in helix_residues.
                atom]
                helix_data.extend(collapse_residue_ranges(helix_residue_numbers))

    cmd.delete("all")
    return helix_data

def process_directory(cif_dir, output_file, threshold=70):
    """Process all CIF files in the given directory and write results to a table."""
    results = []

    for cif_file in os.listdir(cif_dir):
        if cif_file.endswith('.cif'):
            full_cif_path = os.path.join(cif_dir, cif_file)
            print(f"Processing {cif_file}...")
            helix_data = extract_helix_positions(full_cif_path, threshold)

            # Prepare row data
            row = [os.path.basename(cif_file)]
            for start, end in helix_data:
                row.append(start)
                row.append(end)
            results.append(row)

    # Determine the maximum number of helix columns required
    max_helix_count = max((len(row) - 1) // 2 for row in results)

    # Create header
    header = ["Name"]
    for i in range(1, max_helix_count + 1):
        header.append(f"Helix {i} Start")
        header.append(f"Helix {i} End")

    # Write results to a CSV file
    with open(output_file, 'w', newline='') as csvfile:
        writer = csv.writer(csvfile)
        writer.writerow(header)
        for row in results:
            # Pad rows with empty values if they have fewer helices
            while len(row) < len(header):
                row.append("")
            writer.writerow(row)

# Main function to run the processing
def main():
    cif_dir = "cif"  # Directory containing CIF files
    output_file = "results_70.csv"  # Name of output file
    threshold = 70  # pIDDT threshold for selecting residues (B-factor >= 70)

    process_directory(cif_dir, output_file, threshold)

if __name__ == "__main__":
    main()
```

Script 1: Python script for extraction of α-helix positions from AlphaFold protein models via PyMol (Abramson *et al.*, 2024; Schrödinger, 2010).

```python
import os
import csv
import pymol
from pymol import cmd

def collapse_residue_ranges(residues):
    """Collapse a list of domain positions into residue ranges."""
    if not residues:
        return []

    residues = sorted(map(int, residues))  # Convert residues to integers
    collapsed = []

    start = residues[0]
    end = residues[0]

    for i in range(1, len(residues)):
        if residues[i] == end + 1:
            end = residues[i]
        else:
            if start == end:
                collapsed.append((start, end))  # Single residue
            else:
                collapsed.append((start, end))  # Range
            start = residues[i]
            end = residues[i]

    # Append the last range or single residue
    collapsed.append((start, end))
    return collapsed

def extract_sheet_positions(cif_file, threshold=70):
    """Extract beta-sheet positions for chains with pIDDT >= threshold."""
    cmd.load(cif_file)

    sheet_data = []
    for chain in cmd.get_chains():
        cmd.select(f'chain_{chain}', f'chain {chain}')

        atoms = cmd.get_model(f'chain_{chain}')
        selected_residues = [atom.resi for atom in atoms.atom if atom.b >= threshold]

        if selected_residues:
            residue_selection = '+'.join(map(str, selected_residues))
            cmd.select(f'sheet_{chain}', f'chain {chain} and ss S and resi {
            residue_selection}')

            sheet_residues = cmd.get_model(f'sheet_{chain}')
            if sheet_residues:
                sheet_residue_numbers = [int(atom.resi) for atom in sheet_residues.
                atom]
                sheet_data.extend(collapse_residue_ranges(sheet_residue_numbers))

    cmd.delete("all")
    return sheet_data

def process_directory(cif_dir, output_file, threshold=70):
    """Process all CIF files in the given directory and write results to a table."""
    results = []

    for cif_file in os.listdir(cif_dir):
        if cif_file.endswith('.cif'):
            full_cif_path = os.path.join(cif_dir, cif_file)
            print(f"Processing {cif_file}...")
            sheet_data = extract_sheet_positions(full_cif_path, threshold)

            # Prepare row data
            row = [os.path.basename(cif_file)]
            for start, end in sheet_data:
                row.append(start)
                row.append(end)
            results.append(row)

    # Determine the maximum number of sheet columns required
    max_sheet_count = max((len(row) - 1) // 2 for row in results)

    # Create header
    header = ["Name"]
    for i in range(1, max_sheet_count + 1):
        header.append(f"Sheet {i} Start")
        header.append(f"Sheet {i} End")

    # Write results to a CSV file
    with open(output_file, 'w', newline='') as csvfile:
        writer = csv.writer(csvfile)
        writer.writerow(header)
        for row in results:
            # Pad rows with empty values if they have fewer helices
            while len(row) < len(header):
                row.append("")
            writer.writerow(row)

# Main function to run the processing
def main():
    cif_dir = "cif"  # Directory containing CIF files
    output_file = "results_sheet_70.csv"  # Name of output file
    threshold = 70  # pIDDT threshold for selecting residues (B-factor >= 70)

    process_directory(cif_dir, output_file, threshold)

if __name__ == "__main__":
    main()
```

Script 2: Python script for extraction of β-sheet positions from AlphaFold protein models via PyMol (Abramson *et al.*, 2024; Schrödinger, 2010).

```python
import os
import pymol
from pymol import cmd

pymol.finish_launching(["pymol", "-qc"])

def analyze_protein_cif(file_path, output_file, pidt_threshold=70, distance_cutoff=5):

    try:
        # Extract protein names from the filename (format:
        fold_PROTEINA_PROTEINB_model_0.cif)
        structure_name = os.path.splitext(os.path.basename(file_path))[0]
        parts = structure_name.split("_")

        proteinA = parts[1]
        proteinB = parts[2]

        # Load the structure
        cmd.load(file_path, structure_name)

        # Step 1: Select residues with b > pidt threshold for both chains
        cmd.select("chainA_b_factor", f"chain A and b > {pidt_threshold}")
        cmd.select("chainB_b_factor", f"chain B and b > {pidt_threshold}")

        # Step 2: Find residues within distance_cutoff
        cmd.select("close_A", f"chainA_b_factor within {distance_cutoff} of
        chainB_b_factor")
        cmd.select("close_B", f"chainB_b_factor within {distance_cutoff} of
        chainA_b_factor")

        # Step 3: Get the length of each protein
        chainA_residues = set(atom.resi for atom in cmd.get_model("chain A").atom)
        chainB_residues = set(atom.resi for atom in cmd.get_model("chain B").atom)

        # Calculate the number of unique residues in each chain
        chainA_length = len(chainA_residues)
        chainB_length = len(chainB_residues)

        # Extract residue IDs for residues meeting the criteria above
        close_A_residues = sorted(set(int(atom.resi) for atom in cmd.get_model(
        "close_A").atom))
        close_B_residues = sorted(set(int(atom.resi) for atom in cmd.get_model(
        "close_B").atom))

        # Write results to the output file
        with open(output_file, "a") as f:
            f.write(f"Residues in {proteinA} (close to {proteinB}):\n")
            if close_A_residues:
                f.write(", ".join(map(str, close_A_residues)) + "\n")
            else:
                f.write("None\n")

            f.write(f"Residues in {proteinB} (close to {proteinA}):\n")
            if close_B_residues:
                f.write(", ".join(map(str, close_B_residues)) + "\n")
            else:
                f.write("None\n")


        print(f"Analysis complete for {file_path}. Results appended to {output_file}")
    except Exception as e:
        print(f"Error processing {file_path}: {e}")
    finally:
        # Clean up PyMOL objects
        cmd.delete("all")

def analyze_folder(input_folder, output_file, pidt_threshold=50, distance_cutoff=5):

    cif_files = [f for f in os.listdir(input_folder) if f.endswith(".cif")]
    for cif_file in cif_files:
        file_path = os.path.join(input_folder, cif_file)
        analyze_protein_cif(file_path, output_file, pidt_threshold, distance_cutoff)


if __name__ == "__main__":
    # User-specified parameters
    input_folder = "cif"  # Path to the .cif file.
    output_file = "Interacting residues_short.txt"  # Output file
    pidt_threshold = 50  # Minimum pIDDT score
    distance_cutoff = 4.5  # Distance cutoff in Å for residue interaction

    # Run batch analysis
    analyze_folder(input_folder, output_file, pidt_threshold, distance_cutoff)

    # Shutdown PyMOL when done
    cmd.quit()
```

Script 3: Python script for extraction of interacting residues from AlphaFold protein models via PyMol (Abramson *et al.*, 2024; Schrödinger, 2010).

```r
# Load necessary libraries
library(ggplot2)
library(dplyr)
library(patchwork)

##### IMPORTANT ####
# The sequences in "Sequences_R.fa" MUST be sorted like the alphabetically sorted .cif files. For each .cif file, two sequences are present, the one mentioned first as the first one, then the second one (see file names)

##### Intersecting close AS and Protein Sequences ####

# Import and read files
fasta_file <- "Sequences_R.fa" # File containing SORTED fasta sequences
interaction_file_50 <- "Interacting_residues_short_50.txt" # File containing interaction residues in short format (pIDDT50)
interaction_file_70 <- "Interacting_residues_short_70.txt" # File containing pIDDT >70 residues

fasta_lines <- readLines(fasta_file)
interaction_lines_50 <- readLines(interaction_file_50)
interaction_lines_70 <- readLines(interaction_file_70)

# Filter sequences and names from fasta file and removes ">"
sequence_names <- fasta_lines[seq(1, length(fasta_lines), by = 2)]
sequences <- fasta_lines[seq(2, length(fasta_lines), by = 2)]
sequence_names <- gsub("^>", "", sequence_names)

# Filter interaction lines to exclude lines with names (i.e.: Residues in...)
interaction_lines_50 <- interaction_lines_50[!grepl("^Residues", interaction_lines_50)]
interaction_lines_70 <- interaction_lines_70[!grepl("^Residues", interaction_lines_70)]


all_results <- data.frame()

# Process each sequence and its interaction data
for (i in seq_along(sequences)) {
  sequence <- sequences[i]
  sequence_name <- sequence_names[i]

  interaction_positions_50 <- as.numeric(unlist(strsplit(interaction_lines_50[i], ",")))
  interaction_positions_70 <- as.numeric(unlist(strsplit(interaction_lines_70[i], ",")))

  residues <- strsplit(sequence, "")[[1]]
  residue_positions <- 1:length(residues)
  sequence_df <- data.frame(
    Sequence_Name = sequence_name,
    Position = residue_positions,
    Residue = residues
  )

  closeAS_50 <- data.frame(Position_50 = interaction_positions_50)
  closeAS_70 <- data.frame(Position_70 = interaction_positions_70)

  sequence_df <- sequence_df %>%
    mutate(Interacting = if_else(Position %in% closeAS_70$Position, "70",
      if_else(Position %in% closeAS_50$Position, "50", "0")
    ))

  all_results <- bind_rows(all_results, sequence_df)
}

#### Plotting of results: ####
max_length <- max(all_results$Position)
plot_list <- list()

# Loop through each unique sequence name
for (sequence_name in unique(all_results$Sequence_Name)) {
  sequence_data <- all_results %>%
    filter(Sequence_Name == sequence_name)

  p <- ggplot(sequence_data, aes(x = Position, y = 1, fill = Interacting)) +
    geom_tile(height = 1) +
    scale_fill_manual(
      values = c("70" = "#038f0a", "50" = "#19059c", "0" = "#d9d5d4"),
      name = "Residues",
      labels = c("70" = "≤ 4.5 Å & pIDDT > 70", "50" = "≤ 4.5 Å & 50 < pIDDT ≤ 70"
      , "0" = "> 4.5 Å | pIDDT ≤ 50")
    ) +
    scale_x_continuous(limits = c(0, max_length)) +
    labs(
      x = "Residue Position",
      y = sequence_name
    ) +
    theme_minimal() +
    theme(
      aspect.ratio = 0.02,
      axis.text.y = element_blank(),
      axis.ticks.y = element_blank(),
      axis.title.y = element_blank(),
      panel.grid.major = element_blank(),
      panel.grid.minor = element_blank(),
      plot.margin = margin(0, 0, 0, 0)
    )

  plot_list[[sequence_name]] <- p
}

# Remove X-axis labels for all but the last plot (to keep as legend)
plot_list <- lapply(seq_along(plot_list), function(i) {
  if (i < length(plot_list)) {
    plot_list[[i]] + theme(
      axis.title.x = element_blank(),
      axis.text.x = element_blank(),
      axis.ticks.x = element_blank()
    )
  } else {
    plot_list[[i]]
  }
})

# Remove legends for plots without any residues lower than pIDDT 70 or 50. (For combining legend via patchwork)
remove_short_legends <- function(plot, legend_column) {
  # Extract the data used in the legend
  data <- ggplot_build(plot)$data[[1]]

  # Count the number of unique entries in the specified column
  num_unique <- length(unique(data[[legend_column]]))

  if (num_unique < 3) {
    plot <- plot + guides(color = "none", fill = "none")
  }

  return(plot)
}
plot_list <- lapply(plot_list, function(plot) {
  remove_short_legends(plot, "group")
})

# Combine plots in pairs to signify interactions, stack them vertically, and add a combined title
combined_plots <- lapply(seq(1, length(plot_list), by = 2), function(i) {
  plot_pair <- plot_list[c(i, i + 1)]

  plot_pair <- lapply(plot_pair, function(p) {
    p + theme(axis.title.y = element_blank())
  })

  y_titles <- sapply(plot_pair, function(p) {
    title <- ggplot_build(p)$plot$labels$y
    # Remove unwanted suffixes
```

Script 4: R script for visualizing position of extracted interacting residues.

```
        gsub("_p[0-9]+", "", title)
    })
    combined_title <- paste(y_titles[1], "+", y_titles[2])

    # Add title to the first plot of the plot pair
    plot_pair <- lapply(seq_along(plot_pair), function(i) {
        if (i < length(plot_pair)) {
            plot_pair[[i]] + labs(title = combined_title)
        } else {
            plot_pair[[i]]
        }
    })

    combined_plot <- wrap_plots(plot_pair, ncol = 1)
    return(combined_plot)
})

# Combine all paired plots together in a single layout (all stacked vertically)
final_combined_plot <- wrap_plots(combined_plots, ncol = 1) +
    plot_layout(guides = "collect")

ggsave("Interacting_residues.pdf", final_combined_plot, width = 10, height = 15)

```

Script 4 (cont.)

```python
import os
import pymol
from pymol import cmd

# Set the folder containing the protein files and the output directory
input_folder = "E:/HomeOffice_current/Computational/Interacting residues/cif"
output_folder = "E:/HomeOffice_current/Computational/Interacting residues/PyMol visualisation"

os.makedirs(output_folder, exist_ok=True)

def visualize_protein(file_path, output_name):
    # Load the protein complex
    cmd.reinitialize()
    cmd.load(file_path)
    protein_name = os.path.splitext(os.path.basename(file_path))[0]

    # Selection of domains
    cmd.select("AtLUG_LUFS", f"/{protein_name}//A and resi 1-160")
    cmd.select("AtSEU_inter", f"/{protein_name}//B and b>50 within 10 of AtLUG_LUFS")

    # Selection of whole residues where at least one atom is within 4.5 Å and b>50 of the other chain
    cmd.select("AtLUG_AS", "byres AtLUG_LUFS and b>50 within 4.5 of AtSEU_inter and b>50")
    cmd.select("AtSEU_AS", "byres AtSEU_inter and b>50 within 4.5 of AtLUG_LUFS and b>50")

    # Grouping selections
    cmd.group("AtLUG_SEU", f"{protein_name} AtLUG_LUFS AtLUG_AS AtSEU_AS AtSEU_inter")

    # Coloring
    cmd.bg_color("white")
    cmd.color("hafnium", f"/{protein_name}//A")
    cmd.color("actinium", f"/{protein_name}//A and backbone")
    cmd.color("orange", f"/{protein_name}//B")
    cmd.color("copper", f"/{protein_name}//B and backbone")
    cmd.color("grey", f"/{protein_name}/ and b<50")


    # Visualization
    cmd.select("AtSEU_inter", "AtSEU_inter extend 20")
    cmd.hide("everything")
    cmd.show("cartoon", "AtLUG_LUFS AtSEU_inter")
    cmd.show("licorice", "AtLUG_AS AtSEU_AS")

    # Adjust view to fill canvas
    cmd.orient("AtLUG_AS")
    cmd.zoom("AtLUG_AS", buffer=10)
    cmd.set("fog", 0)  # Disables the fog effect

    # Export with desired resolution and orientation
    output_file = os.path.join(output_folder, f"{output_name}.png")
    cmd.ray(4800, 3000)
    cmd.png(output_file, dpi=600)

# Loop through all files in the folder
for filename in os.listdir(input_folder):
    if filename.endswith(".pdb") or filename.endswith(".cif"):
        file_path = os.path.join(input_folder, filename)
        output_name = os.path.splitext(filename)[0]
        visualize_protein(file_path, output_name)

print("Visualization complete. Images saved to:", output_folder)

```

Script 5: Python script for visualizing interactions of the LUFS-domain of LUG with SEU in AlphaFold protein models via PyMol (Abramson *et al.*, 2024; Schrödinger, 2010).

```
//Dynamicaly define input and output folders and file type
#@ File (label = "Input directory", style = "directory") input
#@ File (label = "Output directory", style = "directory") output
#@ String (label = "File type", value = ".jpg") suffix

//Define pre- and suffixe for processed pictures
#@ String (label = "Output prefix", value = "") output_prefix
#@ String (label = "Output suffix", value = "") output_suffix

// function to scan folders/subfolders/files to find files with correct suffix
processFolder(input);
function processFolder(input) {
    list = getFileList(input);
    list = Array.sort(list);
    for (i = 0; i < list.length; i++) {
        if(File.isDirectory(input + File.separator + list[i]))
            processFolder(input + File.separator + list[i]);
        if(endsWith(list[i], suffix))
            processFile(input, output, list[i]);
    }
}

//Execution of macro for each file
function processFile(input, output, file) {


//image handeling and preparation
open(input + File.separator + file);
original_image = getTitle(); //Get image name for later selection
selectWindow(original_image);

//Manual cropping
makeRectangle(1032, 612, 3888, 2580);
waitForUser("Position selection EXACTLY on internal border (edge media - plate wall)
of the plates!");
run("Crop");
// Sometimes, to reduce noise during the analysis, a slight blur is advised
//run("Gaussian Blur...", "sigma=1");

//Saving the image
filename_pure = output_prefix + File.nameWithoutExtension + output_suffix; //Adds
pre- and suffixe to the file name
saving_prefix = output + File.separator + filename_pure; //Defines save location
saveAs("JPEG", saving_prefix);

//Clean-Up for next picture
run("Close All");

}
```

Script 6: Fiji macro for semi-automatic cropping of Y2H assay plate pictures.

```r
## Information & Settings ####
# Goal: Automated Processing (Gitter, rename and analyse) pictures from Y2H-screens

# Use the gitter.batch() function, which allows you to supply a reference screen that is analyzed first and used to create a grid that is overlayed onto the second plate (the plate you want to analyse)
# For this, the images need to be as similar as possible regarding pixel size, position of the plate ect.
# Crop the images to the INSIDE of the plate using an ImageJ script to keep everything consistent
# Check the gitted images visually to confirm that the grids were placed correctly! IMPORTANT!

# Also, after the gitter analysis, check the output images for unwanted artifacts or noise and make sure that all colonies were captured by the analysis
# THIS IS THE ONLY WAY TO ENSURE THAT THE DATA WAS COLLECTED PROPERLY!

# If colonies are not captured correctly, try the following:
# Re-gitter the image. Sometimes this alone gets rid of the noise
# Re-crop the images to make sure your positioning is ok
# Add a small bit of gausian blur (start with 1, slowly increase) in case there is high noise in the gitter_picture plate
# If there are non-captured spots (black) within colonies, they are caused by the general color of the colonies/color differences within the colonies which is wrongfully assumed to be different by the binary transformation
# Try reducing the background in ImageJ using rolling ball substraction


# Gittered reference image that contains all 96 colonies in a good quality
file_ref <- "./Gitter/Plate_reference.jpg"

# Define save location of files
dat_save <- "./Gitter/Output_file/"
grid_save <- "./Gitter/Output_image/"

# Define your source directory
dir_source <- "./Gitter/Source/"

# Gitter reference file
dat_ref <- gitter(file_ref,
   dat.save = dat_save,
   grid.save = grid_save,
   plate.format = 96,
   verbose = "p"
)


# IMPROTANT:
# The script assumes that your files ar named after the following scheme:
# Plate_[ID]v[REPLICATE]_[Treatment]
# [ID] and [REPLICATE] should be numeric values
# [Treatment] should be one of the following
# LW_Transfer [This is the reference plate]
# LWH
# LWH_3-AT


# ANALYSIS:
# Idea: Calculate sizes of LWH and 3-AT colony growth in relation to the LW-transfer colonies
# --> The smaller the difference, the more likely that there is a genuine protein interaction, as this allows for growth of colonies on the selective media


# Definition of selective thresholds
colony_size_threshold <- 900 # Based on a first look, growing spots seem to have a colony size of >1000

# DEFINITION OF WHEN WE CONSIDER COLONY GRWOTH TO BE A GENUINE PROTEIN INTERACTION
# Currently: If assay colonies have grown to a size of 50% or more in relation to control colony
interaction_true <- 0.5


# LOOKUP-TABLE
# To automatically map the row/col coordinates of the plates to the proteins, you must create a lookup table in excel:
# One sheet consisting of seven columns with the following data:
# [Plate], [row], [row_tag], [row_protein], [column], [col_tag], [col_protein]
# [Plate] is plate number WITHOUT replicates (i.e 1, 2, 3 ect.)
# [row/col] is position of row/col
# [tag] is AD/BD
# [protein] is protein name

# Load lookup-table
library(readxl)
lookup_grid <- read_excel("./Y2H_Lookup_table.xlsx", sheet = "grid")

# Merge the tag and protein columns to one combined column and remove the old ones
lookup_grid <- lookup_grid %>%
   mutate(tag_prot_row = paste0(tag1, "_", protein1)) %>%
   select(!c(tag1, protein1))

lookup_grid <- lookup_grid %>%
   mutate(tag_prot_col = paste0(tag2, "_", protein2)) %>%
   select(!c(tag2, protein2))

# Change plate number from "Double" to "Character" and add "Plate_" for later merge with results
lookup_grid$plate <- as.character(lookup_grid$plate)
lookup_grid$plate <- paste0("Plate_", lookup_grid$plate)


## Setup ####

# Install necessary packages
install.packages("gplots")
install.packages("LSD")
install.packages("RColorBrewer")
install.packages("plyr")
install.packages("stringr")
install.packages("logging")

install.packages("devtools")
library(devtools)
devtools::install_github("omarwagih/gitter")

install.packages("BiocManager")
library(BiocManager)
BiocManager::install("EBImage")
BiocManager::install("ComplexHeatmap")

# Load necessary packages

library(gitter)
library(dplyr)
library(tidyr)
library(RColorBrewer)
library(ggplot2)
library(writexl)

## Batch Gitter ####
gitter.batch(dir_source,
   ref.image.file = file_ref,
   dat.save = dat_save,
   grid.save = grid_save,
   plate.format = 96,
   verbose = "p"
)


# This code loads multiple tab-separated gitter data files, appends them row-wise, and adds columns indicating the source file and treatment for each row:
combined_gitter_data <- function(directory) {
   gitter_data_list <- list.files(dat_save, pattern = "\\.dat$", full.names = TRUE)
   gitter_data_inter <- list()
```

Script 7: R script using the gitter package to measure yeast colony sizes and to determine presence/absence of protein interactions in the assay (Wagih & Parts, 2014).

```r

    # The col names are unfortunatly marked as comments, so we need to supply them manually
    colnames_gitter <- c("row", "col", "size", "circularity", "flags")

    for (file in gitter_data_list) {
       data_f <- read.table(file,
          comment.char = "#",
          sep = "\t"
       )
       colnames(data_f) <- colnames_gitter

       # Splits the file name into separet stings at each "_" in the name to precisely name the plates and treatments
       split_name <- strsplit(basename(file), "_")[[1]]
       data_f$plate <- paste(split_name[1], split_name[2], sep = "_")

       # An if loop that creates a treatment column and checks if the fourth string starts with "cropped"
       # Will be the case for LWH-plates, but not LW_Transfer or LWH_3AT
       if (length(split_name) >= 4) {
          if (startsWith(split_name[4], "cropped")) {
             data_f$treatment <- paste(split_name[3])
          } else {
             data_f$treatment <- paste(split_name[3], split_name[4], sep = "_")
          }
       } else {
          data_f$treatment <- paste("Unknown")
       }
       gitter_data_inter[[file]] <- data_f
    }
    gitter_data_inter <- do.call(rbind, gitter_data_inter)
    rownames(gitter_data_inter) <- 1:nrow(gitter_data_inter)
    return(gitter_data_inter)
}
gitter_data <- combined_gitter_data()


## Analysis of results ####

# Sets the names of the treatments to numerical values, Easier for referencing and working with them
gitter_data <- gitter_data %>% mutate(
    treatment =
       recode(treatment,
          "LW_Transfer" = 1,
          "LWH" = 2,
          "LWH_3AT" = 3,
          "Unknown" = 4
       )
)
gitter_data <- filter(gitter_data, treatment != 4)


#### Function to obtain and process data from singular plates ####
process_plate <- function(gitter_data) {
    #### #### Extracting data from singular plates and creating a three-dimensional matrix of position and treatment ####

    unique_rows <- sort(unique(gitter_data$row))
    unique_cols <- sort(unique(gitter_data$col))
    unique_treatments <- sort(unique(gitter_data$treatment))
    plate_matrix <- array(NA,
       dim = c(length(unique_rows), length(unique_cols), length(unique_treatments)),
       dimnames = list(unique_rows, unique_cols, unique_treatments)
    )

    for (treat in unique_treatments) {
       treatment_data <- gitter_data %>% filter(treatment == !!treat)
       for (i in 1:nrow(treatment_data)) {
          row_index <- which(unique_rows == treatment_data$row[i])
          col_index <- which(unique_cols == treatment_data$col[i])
          plate_matrix[row_index, col_index, treat] <- treatment_data$size[i]
       }
    }

    #### #### The analysis that is being done with the data ####

    # Sets the LW-Media as the base for the next analysis steps
    base_treatment <- unique_treatments[1]

    # Create a mask for none-growing colonies on the LW-transfer treatment (1)
    threshold_mask <- abs(plate_matrix[, , base_treatment]) < colony_size_threshold

    treatment_matrix <- array(NA, dim = dim(plate_matrix), dimnames = dimnames(plate_matrix))
    for (treat in unique_treatments) {
       if (treat != base_treatment) {
          treatment_matrix[, , treat] <- plate_matrix[, , treat] / plate_matrix[, , base_treatment]
          treatment_matrix[treatment_matrix > 1] <- 1
       }
    }

    for (treat in unique_treatments) {
       if (treat != base_treatment) {
          treatment_matrix[, , treat][threshold_mask] <- NA
       }
    }
    treatment_df <- as.data.frame(as.table(treatment_matrix))
    colnames(treatment_df) <- c("row", "col", "treatment", "rel.size")

    interaction_treat <- treatment_df %>%
       filter(rel.size >= interaction_true)

    list(
       "Gitter analysis" = plate_matrix,
       "Assay effects" = treatment_matrix,
       "No colony growth" = threshold_mask,
       "True interactions" = interaction_treat
    )
}


#### Function to loop the analysis through all plates ####
analyze_all_plates <- function(gitter_data) {
    plate_list <- split(gitter_data, gitter_data$plate)
    plate_list <- plate_list[sapply(plate_list, function(gitter_data) {
       any(gitter_data$treatment == "1")
    })]
    gitter_results <- list()

    # Process each plate by looping through each plate via their ID
    for (plate_id in names(plate_list)) {
       cat("Processing plate", plate_id, "\n")
       plate_analysis <- plate_list[[plate_id]]
       gitter_results[[plate_id]] <- process_plate(plate_analysis)
    }
    gitter_results
}
results <- analyze_all_plates(gitter_data)


#### Function to obtain a dataframe of all true interactions ####
get_protein_interaction <- function(results) {
    interaction_results <- data.frame()
    for (plate_id in names(results)) {
       plate <- results[[plate_id]]
       if ("True interactions" %in% names(plate)) {
          interaction_df <- plate$`True interactions`
       }
       if (nrow(interaction_df) == 0) {
          next
       }

       interaction_df$plate <- plate_id
```

Script 7 (cont.)

```
        interaction_results <- bind_rows(interaction_results, interaction_df)
    }

    # Modify the interaction_results table to store the replicate number in a
    different column to allow for merging of the results with the lookup table names

    interaction_results <- interaction_results %>%
        mutate(replicate = gsub("Plate_[0-9](.*)", "\\1", plate)) %>%
        mutate(plate = gsub("v[0-9]+$", "\\1", plate))

    interaction_results$row <- as.numeric(interaction_results$row)
    interaction_results$col <- as.numeric(interaction_results$col)

    merged_results <- interaction_results %>%
        inner_join(final_lookup, by = c("plate", "row", "col")) %>%
        mutate(treatment = case_when(
            treatment == "2" ~ "LWH",
            treatment == "3" ~ "LWH_3-AT",
            TRUE ~ treatment
        )) %>%
        separate(tag_prot_row, into = c("tag_row", "prot_row"), sep = "_", extra =
        "merge") %>%
        separate(tag_prot_col, into = c("tag_col", "prot_col"), sep = "_", extra =
        "merge")
    return(merged_results)
}
Assay_results <- get_protein_interaction(results)


# Shorten results: How many times (i.e replicates and both directions of
interactions) is a interaction present in the Assay_results

short_and_sort <- function(Assay_results) {
    Assay_results_empty <- Assay_results %>%
        filter(prot_row == "empty" | prot_col == "empty") %>%
        group_by(tag_row, prot_row, tag_col, prot_col) %>%
        summarize(count = n(), .groups = "drop")

    Assay_results_M <- Assay_results %>%
        mutate(tag_prot_row = paste0(tag_row, "_", prot_row)) %>%
        select(!c(tag_row, prot_row))

    Assay_results_M <- Assay_results_M %>%
        mutate(tag_prot_col = paste0(tag_col, "_", prot_col)) %>%
        select(!c(tag_col, prot_col))


    Assay_results_short <- Assay_results_M %>%
        rowwise() %>%
        mutate(
            Sorted_Names = list(c(treatment, sort(c(tag_prot_row, tag_prot_col))))
        ) %>%
        ungroup() %>%
        mutate(
            treatment = sapply(Sorted_Names, `[`, 1),
            protein_1 = sapply(Sorted_Names, `[`, 2),
            protein_2 = sapply(Sorted_Names, `[`, 3)
        ) %>%
        group_by(treatment, protein_1, protein_2) %>%
        summarize(
            count = n(), .groups = "drop"
        ) %>%
        separate(protein_1, into = c("tag_1", "prot_1"), sep = "_", extra = "merge") %>%
        separate(protein_2, into = c("tag_2", "prot_2"), sep = "_", extra = "merge")

    write_xlsx(
        list(
            Assay_results = Assay_results,
            Empty_interactions = Assay_results_empty,
            Assay_results_short = Assay_results_short
        ),
        "./Assay_results.xlsx"
    )
```

```
}

short_and_sort(Assay_results)


#### Visualizing Analysis Output ####

generate_bubble <- function(results) {
    # Recode the treatment names for better overview (for use in plot only!)
    treatment_labels <- c(
        "1" = "LW",
        "2" = "LWH",
        "3" = "LWH + 3-AT"
    )
    pdf(
        file = "./Y2H_results_plots.pdf",
        width = 8, height = 6
    )

    for (plate_id in names(results)) {
        plate_results <- results[[plate_id]]
        treatment_results <- plate_results$`Assay effects`
        treatment_name <- dimnames(treatment_results)[[3]]
        for (treatment in treatment_name) {
            # Remove analysis of LW-plate against LW-plate from dataframe (treatment =
            1), as the output is just NAs
            bubble_matrix <- treatment_results[, , treatment]
            if (all(is.na(bubble_matrix))) {
                next
            }

            bubble_df <- as.data.frame(as.table(bubble_matrix))
            colnames(bubble_df) <- c("row", "col", "rel.size")


            # Replace NA values for missing colonies with -1, a value that is otherwise
            impossible
            bubble_df <- bubble_df %>%
                mutate(rel.size = if_else(is.na(rel.size), -1, rel.size))

            bubble_df <- bubble_df %>%
                mutate(threshold = if_else(rel.size < interaction_true, "FALSE", "TRUE"))

            print(
                plot <- ggplot(
                    bubble_df,
                    aes(x = col, y = row)
                ) +
                    geom_point(shape = 21, fill = "darkgrey") +
                    geom_point(
                        shape = 21,
                        aes(size = rel.size, fill = threshold)
                    ) +
                    scale_radius(
                        range = c(3, 15),
                        limits = c(0, 1),
                        breaks = c(0, 0.25, 0.5, 0.75, 1),
                        labels = c("NA", "0.25", "0.5", "0.75", "1")
                    ) +
                    scale_fill_manual(
                        values = c(
                            "FALSE" = "#F61E1E",
                            "TRUE" = "#4ADB2C"
                        ),
                        name = "Protein interaction"
                    ) +
                    labs(
                        title = paste("Bubble Plot for Treatment:", plate_id, ",",
                        treatment_labels[treatment]),
                        x = "Column",
                        y = "Row"
                    ) +
                    guides(
```

Script 7 (cont.)

```
                    fill = "none",
                    size = guide_legend(
                        title = "Rel. size",
                        override.aes = list(
                            size = c(1.5, 6, 9, 12, 15),
                            fill = c("darkgrey", "#F61E1E", "#4ADB2C", "#4ADB2C",
                            "#4ADB2C")
                        )
                    )
                ) +
                scale_y_discrete(limits = rev) +
                theme_minimal() +
                theme(
                    legend.position = "right",
                    panel.grid.major = element_blank(),
                    panel.border = element_rect(fill = NA),
                    axis.ticks = element_line(color = "black")
                ) +
                coord_fixed()
            )

            cat("Processing plate", plate_id, ",", treatment_labels[treatment], "\n")
        }
    }
    dev.off()
}

generate_bubble(results)

```

Script 7 (cont.)

```
//Dynamicaly define input files, output folders and file type
#@ File (label = "DAPI file", style = "open") image_DAPI
#@ File (label = "YFP file", style = "open") image_YFP
#@ File (label = "Output directory", style = "directory") output
#@ String (label = "File type", value = ".tif") suffix

//Define pre- and suffixe for processed pictures
#@ String (label = "Merge suffix", value = "_merge") output_sufMerge
#@ String (label = "Crop suffix", value = "") output_sufCrop


//Open DAPI-Picture
open(image_DAPI);
SaveDAPI = File.nameWithoutExtension //Stores the Name of the file for saving later
Image_DAPI = getTitle(); //Allows for selection of the opened file
run("Duplicate...", "title=Duplicate");

//Split channels, keep only the blue one
run("Split Channels");
selectImage("Duplicate (green)");
close();
selectImage("Duplicate (red)");
close();


//Open YFP-Picture
open(image_YFP);
SaveYFP = File.nameWithoutExtension //Stores the Name of the file for saving later
Image_YFP = getTitle() //Allows for selection of the opened file
selectImage(Image_YFP);
run("Duplicate...", "title=Duplicate2");

//Split channels, keep only the green one
run("Split Channels");
selectImage("Duplicate2 (blue)");
close();
selectImage("Duplicate2 (red)");
close();


//Merging of blue DAPI and green YFP channels
run("Merge Channels...", "c2=[Duplicate2 (green)] c3=[Duplicate (blue)]");
selectImage("RGB");

//Add selection for cropping
run("Specify...", "width=750 height=750 x=1 y=1");
waitForUser("Position ROI, then hit OK");
//You need to select the area to crop manually. Select a well visible/overlayed area
for the cropping


//Crop and save Merged image
roiManager("Add");
run("Crop");
saveAs("PNG", output + File.separator + File.nameWithoutExtension + output_sufMerge);

//Crop and save YFP and DAPI images
selectImage(Image_DAPI);
roiManager("Select", 0);
run("Crop");
saveAs("PNG", output + File.separator + SaveDAPI + output_sufCrop);
selectImage(Image_YFP);
roiManager("Select", 0);
run("Crop");
saveAs("PNG", output + File.separator + SaveYFP + output_sufCrop);


//Cleanup for next pixture
roiManager("Select", 0);
roiManager("Deselect");
roiManager("Delete");
run("Close All");
```

Script 8: Fiji macro to split and merge color channels for BiFC analysis.

Note of Thanks

To finish this dissertation, I want to thank multiple people without whom this work would have scarcely been possible:

For one, I want to sincerely thank my colleagues that were critical in the creation of several pieces of this dissertation: Especially Andrea Weisert, who performed a brunt of the cloning work in regards to the Y2H and BiFC assay vectors, and whose knowledge about a variety of wet lab work tremendously helped in troubleshooting and streamlining the various experimental procedures I had to perform, but also Clemens Rössner, who not only performed some of the analyses presented in this work, but also contributed several thoughts and improvements to various aspects of this thesis and our shared paper (even though we probably should have talked more about our work than we usually did...). Additionally, I want to thank my colleagues from the lab of Prof. Stefanie Müller-Schüssele, especially Julian Ingelfinger and Aurora Martins, who were vital for the transformation of *P. patens* and created the fascinating mutant lines described here, as well as Farima Moafi, who helped me in genotyping the hundreds of available mutant lines and managed to identify the *pplug3 pplug4* lines I presented and analyzed here. Similarly, Vishnu Narayanan Suma Sreechakram of the lab of Prof. Sigurd Braun helped me a lot in performing the robot-assisted Y2H assays, without which this large-style assay would not have been possible.

Secondly, I want to thank all members of the Becker lab, for creating a warm and helpful working environment, and who made my time there a nice and productive experience. Especially Siwei Pang, Doudou Kong, Le-Han Rössner, and Clemens Rössner helped me to stay reasonably sane during my time there, with lively discussions about our work, and more importantly, about various non-work related stuff. I also want to thank Katheryn Pisfil Colchado, Katrin Ehlers, Claudia Jung-Blasini, and several former members, like Linus Wegner, Dominik Lotz, and Romain Scalone, who helped me with and in a variety of tasks, thoughts, and ideas over the last several years. I also want to thank Prof. Annette Becker for her continuous support and supervision of my work, for her ideas and input regarding problems or further experiments, and for giving me this opportunity in the first place.

My thanks also extend to the various professors, doctors, and especially PhD students of the ICIPS research unit, who contributed to this dissertation and my work at large, and from whom I learned a lot about different aspects and intricacies of plant reproduction, about fascinating new methods, and about a variety of science-adjacent things in general. Especially the ability to communicate and exchange ideas, methods, and data, and to work closely together with different labs interested in the same topic as I was an extremely productive experience and helped to shape several aspects of my various projects.

I also want to thank the DFG (German Research Foundation) for financing my PhD position, Prof. John Golz for his collaboration with our paper and for several plasmids used in

the Y2H assay, as well as Prof. Richard Imminck and Florian Rümpler for sharing various vectors used.

Similarly, I also want to thank the members of my PhD defense committee, Prof. Annette Becker, Prof. Günter Theißen, Prof. Michael Frei, and Prof. Sigurd Braun, for their time and effort during this process.

Finally, I also want to thank my friends and especially my family for supporting my aspirations and for listening to me venting about problems I encountered, but also about my successes; and being available for me during these last few years, be it by moving small spaceships around a table, playing improv theater, shooting a series (where we definitely didn't overdo it, scope-wise...), or simply just spending time together.

Affidavit

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Gießen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.

Statement about the use of Artificial Intelligence (AI) based aids like ChatGPT or SchulKI by OpenAI, or Gemini by Google in the creation of my thesis (marked as applicable):

☐ I have not used any Al tool in preparing this text

■ I used an Al tool in the following areas (multiple answers possible):

☐ Finding ideas, stimulating my creativity

■ Understanding concepts, researching facts and definitions

☐ Optimizing a text that I drafted myself

☐ Creating entire text passages following my prompts

I used the following AI tools to improve the given passages of the text in the manner stated:

AI tools were used to help with programming-related tasks for the scripts included in the supplementals, by helping with the creation and troubleshooting of these scripts. Additionally, AI was used to help tweak both the output of graphs created in R, and the layout of this thesis in LaTeX. These tweaks were of optical nature only; no content, text, or data was created or modified by AI. AI tools were also used to search for specific topics in published literature, but were not used to summarize or cite used literature, nor to write any passages of this dissertation.

Date: _____ Signature: _____