Title: Evolution and protein interactions of LEUNIG and SEUSS homologs across land plants

Authors: Julian Garrecht[1], Clemens Rössner[1], Vishnu Narayanan Suma Sreechakram[2], Qiwei Li[3], Sigurd Braun[2], John Golz[3], Annette Becker[1]

Affiliations

1 Institute of Botany, Justus-Liebig-University Gießen, Germany

2 Institute of Genetics, Justus-Liebig-University Gießen, Germany

3 School of BioSciences, The University of Melbourne, Melbourne, Australia

Abstract

The evolution of land plants has involved significant restructuring and expansion of gene networks responsible for developmental processes, leading to the emergence of new expression patterns and gene activities. The transcriptional co-regulators LUG (LEUNIG) and SEU (SEUSS) play crucial roles in *Arabidopsis thaliana*'s sexual reproduction, participating in various development processes required for forming angiosperm-specific features. While *LUG* and *SEU* have existed for at least 500 million years, their exact phylogenetic relationship, and when they became protein interaction partners remain unclear. We carried out phylogeny reconstruction, protein domain analysis and comparative transcriptome analysis to elucidate the evolutionary dynamics of the LUG and SEU gene homologs across land plants, revealing insights into protein interactions and lineage-specific adaptations. LUG and SEU proteins from diverse land plant lineages interact via the same protein regions and these domains are also found in Zygnematophyceae, suggesting that LUG-SEU dimerization predates land plants but is not found in other Streptophyte algae lineages. Our findings suggest that while physical interactions are conserved among the LUG/SEU proteins in land plants and beyond, there exist lineage-specific differences in expression patterns and domain organization, which may contribute to functional diversification of LUG and SEU during evolution. Further research is warranted to elucidate the structural and functional implications of these variations across diverse plant lineages.

Introduction

During land plant evolution, gene networks governing developmental processes rewired and expanded tremendously, allowing for the generation of novel expression patterns and the regulation of novel genes. Gene duplications, changes in gene expression and protein domains structure of transcription factors and co-regulators are drivers of regulatory diversity. The origin of novel reproductive organs, such as the land plant-specific antheridia and archegonia, pollen, stamens, and ovules that emerged in the lineage leading to seed plants and the angiosperm-specific carpels involved the evolution of gene regulatory networks specifying organ and tissue identity for their initiation and development (Becker et al. 2025).

Unlike many transcription factors that act specifically in the development of one to few tissues or organs in plants, the eukaryotic transcriptional co-regulators LEUNIG and SEUSS are involved in the development of all reproductive organs in *Arabidopsis thaliana* via binding to a multitude of transcription factors of different families. LUG-like genes (LUG and

LEUNIG_HOMOLOG (LUH) in Arabidopsis) encode transcriptional regulators in plants, with similarities to proteins in yeast, fruit flies and humans, belonging of the Gro/TLE super family of transcriptional co-repressors (Liu und Meyerowitz 1995; Conner und Liu 2000; Liu und Karmarkar 2008). The LUG protein family members contain a LUFS domain (Sridhar et al. 2004; Liu und Karmarkar 2008) near its N-terminal end, a central glutamine-rich region, and several WD repeats at their C terminus. The LUFS domain mediates interaction with members of the SEU protein family. LUG-mediated target gene repression is largely based on LUG's interaction with the class A histone deacetylase HDA19, a histone modifying enzyme (Gonzalez et al. 2007). A smaller number of targets are repressed by LUG's interaction with HEN3 and Med14, members of the mediator complex that enables direct contact to and regulation of RNA Pol II activity (Gonzalez et al. 2007). Thus, LUG interaction with SEU, histone modifiers/mediator complex members, and transcription factors (TFs) leads to the repression of the TF targets, for most targets by modifying chromatin structure, and for a few targets by direct interference with RNA Pol II.

SEU and the related three *SEUSS-like (SLK1-3)* genes of Arabidopsis encode a small group of transcriptional adaptors in plants that include glutamine-rich regions amd a conserved domain with sequence similarity to the LIM-domain-binding (Ldb) family of TFs, described in animals and yeasts (Franks et al. 2002; Bao et al. 2010). SEU proteins lack, as do other Ldb-containing proteins, a recognizable DNA-binding domain and require DNA-binding cofactors to regulate target gene expression. SEU-like proteins are thought to act as an adaptor protein attaching the transcriptional repressors of the LUG protein family to target TFs (Sridhar et al. 2006), but they may also act without LUG. For root stem-cell fate determination in Arabidopsis, SEU is required to assembles a transcriptional complex by physically interacting with the TF SCARECROW to promote the expression of WOX5, the main root stem cell organizing TF At the WOX5 promoter, SEU recruits a SET DOMAIN GROUP methyl transferase (SDG4, also named ASHR3) which induces H3K4me3 leading to activation of WOX5 expression. SEU interacts with SCR via its LBD domain, whereas interactions with the histone methyl transferase involves the N-terminal Q-rich domain (Zhai et al. 2020). Direct protein interaction of SEU was also demonstrated for several target transcription factors involved in reproductive development, such as AP1, ARF3, ARR14, BP, KNAT, SEP3, STM, WUS, YAB1, YAB3 (Franks et al. 2002; Bao et al. 2010; Pfluger und Zambryski 2004; Herrera-Ubaldo et al. 2023)). However, some TFs can bind to LUG alone, for example, IND, YAB1, YAB3, FIL, JAG, and PHV, but most of these proteins also interact with SEU. At present, it is largely unclear which domain of LUG and SEU mediates interaction with the TFs of different TF families. Moreover, *LUG* and *LUH* act partially redundantly, as do the *SEU* and *SLK1-3* genes (Bao et al. 2010; Sitaraman et al. 2008).

Drought stress causes severe delay of reproductive organ development in Arabidopsis lowering reproductive success and of many genes involved in stamen and pistil development show reduced expression under drought stress conditions. However, this developmental delay or even arrest is reversible when drought stress is relieved (Su et al. 2013), but the molecular mechanism of concerted down-regulation of the developmental regulators remains unclear. SEU was reported to play a major role in osmotic stress response, which is imposed by drought or salinity stress. SEU is an essential regulator for transcriptional activation of several major drought-stress responsive genes in Arabidopsis (Wang et al. 2022). Unexpectedly, this regulatory process occurs either independently or upstream of abscisic acid (ABA). Further, the work by Wang et al., (2022) demonstrates that SEU proteins form liquid-like nuclear condensates by conformational changes to the N-terminal intrinsically disordered region

(IDR1), such that SEU adopts a more compact state upon increase of extracellular osmolarity (Wang et al., 2022). Further, *SEU* expression, unlike *LUG* or *LUH* expression is highly upregulated in high sucrose concentrations (Sitaraman et al. 2008; Bao et al. 2010). Thus, SEU may act as a nuclear sensor for osmotic stress (Wang et al., 2022), linking drought stress response to developmental regulation, possibly via chromatin modifications. Interestingly, SEU homologs from dicots and monocots, and the liverwort *Marchantia polymorpha* form condensates in high salinity conditions. *M. polymorpha* condensates are found throughout the cytoplasm, while in Arabidopsis and most other angiosperms, SEU condensates are nuclear localized (Wang et al., 2022).

Reproductive development under drought stress requires tight coordination of drought-stress responsive genes and those genes that direct reproductive development. LUG and SEU are good candidates to fine-tune this coordinated gene expression regulation, because in *A. thaliana*, they act in several developmental processes essential for sexual reproduction. (1) LUG/SEU + APETALA1 (AP1) repress the floral homeotic gene AGAMOUS, thereby providing temporal and spatial clues for reproductive organ initiation in the floral center (Gregis et al. 2006; Gregis et al. 2009). (2) LUG + EAR motif-containing adapter protein (ECAP) are required for microspore development in the anthers (Shi et al. 2024). (3) SEU + unknown proteins are important for megagametophyte development (Bao et al. 2010). (4) LUG/SEU + INNER NO OUTER (INO) enable formation of the outer ovule integument (Simon et al. 2017). (5) LUG + AINTEGUMENTA (ANT) promote the carpel marginal meristem development resulting in ovule initiation and carpel fusion (Liu et al. 2000). Taken together, the LUG and SEU are involved in the development of several angiosperm-specific traits such as carpel marginal tissue formation and the specification reproductive organ location in the flower. But it is also essential for micro- and megagametophyte development in angiosperms, and expression of the LUG/SEU module in gametophytes and sporophytes was shown in the bryophyte *Anthoceros agrestis* and the fern *Ceratopteris richardii* (C-fern) (Li et al. 2020; Marchant et al. 2022).

Previous phylogeny reconstructions have shown that LUG and SEU gene families date back to at least the Bryophyte lineage (Pfannebecker et al. 2017), but they were done before high quality genome sequence was available for streptophyte algae and representatives of all land plant lineages (Bryophytes, Monilophytes, seed plants), leaving gaps on the duplication history of these gene families. Thus, LUG and SEU homologs were present during at least 500 million years of land plant evolution, however, we lack information on their phylogenetic context and when they became protein interaction partners. Their ability to interact allows fine-tuning of target gene expression, and this may have been a prerequisite for the emergence and evolution of the complex body plans of land plants.

Here, we provide a thorough phylogenetic analysis of LUG and SEU homologs from all major land plant lineages and streptophyte algae representatives, documenting an overall moderate number of genes in the analyzed plant lineages and only few duplications giving rise to subfamilies. This is unexpected, because transcriptional regulators tend to maintained after whole genome duplications (WGDs) at a higher proportion than other genes. Land plants underwent multiple rounds of WGDs with many lineage-specific WGDs, often leading to large families of transcriptional regulators. Our domain structure analysis corroborates the observation that these two gene families show little change over the past 500 million years. Further, protein interaction analysis between LUG and SEU homologs of the major land plant lineages suggest that dimer formation already in the last common ancestor of all land plants,

rendering the LUG and SEU protein families as prime examples for coevolution at the scale of proteins.

Materials and Methods

Phylogeny reconstructions and protein domain/structure analyses

Sequences of full length SEU- and LUG-like proteins of *Arabidopsis thaliana, Arabidopsis lyrata, Cleome violacea, Carica papaya, Phaseolus vulgaris, Cicer arietinum, Lens culinaris, Vicia faba, Vitis vinifera, Lepidum sativum, Helianthus annuus, Solanum tuberosum, Solanum lycopersicum, Aquilegia coerulea, Eschscholzia californica, Oryza sativa, Zea mays, Nymphaea colorata, Amborella trichopoda, Thuja plicata, Cycas panzhihuaensis, Ginko biloba, Adiantum capillus-veneris, Azolla filiculoides, Salvinia cucullata, Ceratopteis richardii, Selaginella moellendorffii, Diphasiastrum complanatum, Physcomitrium patens, Marchantia polymorpha, Anthoceros agrestis, Sphagnum fallax, Sphagnum magellanicum, Ceraton purpureus, Mesostigma viride, Klebsormidium nitens, Chlorokybus atmophyticus, Penium margaritaceum*, and *Chara braun*i were acquired using BLAST-search (Altschul et al. 1990) from Phytozome (https://phytozome-next.jgi.doe.gov/) or the ICIPS-garden BLAST server (http://134.176.27.173/blast/, Roessner et al. 2024). The sequences were aligned using MAFFT (Katoh et al. 2002). Maximum-likelihood phylogenies were generated using IQ-TREE2 (Minh et al. 2020) with 100 bootstraps. All computing was carried out using de.NBI VM large (28 VCPUs, 64 GB RAM). Sequences of ingroup SEU- and LUG-like proteins were fed into NCBI conserved domains (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, (Wang et al. 2023). Positions of putative alpha helices were identified using AlphaFold (Abramson et al. 2024). Positions of alpha-helices, Lim binding-, WD-, LisH-domains and positions Glutamin residues were visualized using R 4.2.3.. Protein structures and interactions were predicted using AlphaFold 3 (Abramson et al. 2024). The computed structural models were visualized and modified using open source PyMol 3.0 (Schrödinger LLC 2010). A python script was used to extract the position of α-helices with a pIDDT of ≥ 70 from the modeled SEU orthologs via PyMol (Supplemental script 1).

For analysis of interacting AA residues, closeness of residues was defined as a residue with a pIDDT confidence of > 50 within 4.5 Å of a residue with a pIDDT confidence of > 50 of the other protein chain (Abramson et al. 2024; Parvathy et al. 2024). Two python scripts were used to a) visualize the first 160 AA of the LUG protein as well interacting SEU residues and 20 atoms around those regions (Supplemental script 2), and b) to extract the position of all interacting residues which were visualized in R version 4.4.1 (R Core Team 2022) and RStudio (Posit team 2024) using using ggplot2 (Wickham 2016) (Supplemental scripts 3 and 4).

Digital gene expression analysis

Transcriptome data were obtained from (Mergner et al. 2020) for *Arabidopsis*, from (Lang et al. 2018; Sreedasyam et al. 2023) for *P. patens*, from (Marchant et al. 2022) for *C. richardii* and from (Briginshaw et al. 2022; Bowman et al. 2017; Frank und Scanlon 2015; Higo et al. 2016; Hisanaga et al. 2021; Julca et al. 2021) for *M. polymorpha and* visualized using the pheatmap (Kolde 2019) and ComplexHeatmap (Gu 2022) packages for R.

Results and discussion

LEUNIG and SEUSS homologs were present in the last common ancestor of land plants

To elucidate the evolutionary relationships of *LUG* and *SEU* homologs, phylogeny reconstructions and protein sequence analyses were carried out with LUG and SEU homologs of representatives of all major land plant lineages (Bryophytes and vascular plants including Lycophytes, Monilophytes, gymnosperms and angiosperms) as well as Streptophyte algae, using 39 species with reference quality genomes. Fig. 1 and Supplemental Figure 1 show that SEU homologs are present in all genomes examined in this study, with each homolog having a characteristic LIM-binding domain, with an adjacent glutamine-rich region. Furthermore, structural predictions identify at least two alpha helices in N-terminal of the LIM-binding domain of all SEU homologs. Phylogenetic analysis indicates that SEU homologs fall into three monophyletic clades, the SEU-clade, the SLK-clade and the preSEU/SLK clade. The SEU-clade includes representatives of all major land plant lineages, whereas the SLK-clade includes only sequences derived from vascular plants. The third clade, termed pre-SEU/SLK clade includes sequences of streptophyte algae, bryophytes, lycophytes, monilophytes and gymnosperms but lacks angiosperm sequences. The monophyly of the SEU- and SLK-clades has robust support (94% bootstrap), with the SEU- and SLK-clade each having bootstrap values of 72% and 49%, respectively. The pre-SEU/SLK clade has a 99% bootstrap value. Our findings suggest a loss of pre-SEU/SLK clade genes in angiosperms and SLK clade genes in bryophytes, lycophytes and monilophytes. Interestingly, the lycophyte *Selaginella moellendorfii* (spikemoss, lycophyte) lost the SEU-clade member, but this may be a species-specific loss.

Analogous to SEU members, LUG homologs are also present in all land plants and in streptophyte algae (Fig. 2 and Supplemental Fig. 2). All homologs identified in this analysis include two to five WD domains at the C-terminal region of the protein and several adjacent Q-rich stretches. Most land plant sequences include a LisH domain, which is required for SEU protein binding. However, some land plant proteins that lack the LisH domain may not be annotated properly, as their N-terminal protein region is short when compared to the LisH containing proteins. Only a single streptophyte algae sequence (from *Klebsormidium nitens*) includes the LisH domain but the other algae sequences may also not be annotated correctly, with the first exon possibly missing. Our phylogeny shows that the land plant LUG and LUH proteins are sister to the LUG/LUH sequences present in streptophyte algae with 100% bootstrap support. However, only seedplant LUG and LUH sequences are supported as monophyletic clades (83% and 85% bootstrap support, respectively) while monilophytes LUG homologs are sister to seedplant LUG homologs with 74% bootstrap support. Other nodes showed a bootstrap support <60% and were considered as an unresolved polytomy harboring bryophyte and lycophyte sequences as well as both monophyletic clades.

In summary, our phylogeny reconstructions identifies LUG and SEU homologs in all land plant lineages and Streptophyte algae with several subfamilies well supported, reporting novel LUG and SEU homologs from Streptophyte algae. However, in both phylogenies, algae sequences constitute subfamilies distinct from those of land plants.

LUG and SEU homologs show high domain structure conservation

Domain analyses and protein folding predictions across streptophyte algae and land plant LUG and SEU homologs support this phylogeny in that the larger domains and their positions are conserved in the majority of streptophyte algae and land plant sequences (Fig. 1 and 2). The

canonical Arabidopsis SEU and SLK1-3 proteins have an N-terminal region of around 180 to 300 amino acids (aa), followed by a LIM-binding domain of around 300 aa length and a C-terminal region covering between 220 to 290 aa. While Q-rich stretches are present throughout the N-terminal and C-terminal regions, most of them are found in the N-terminal region close to the LIM-binding domain. Interestingly, we find conserved positioning of the LIM-binding domain and overall lengths of SEU proteins only in the SEU clade, with the exception of the single *Anthoceros agrestis* (hornwort) representative and one of the *Zea mays* (corn) proteins. In the pre-SEU/SLK clade, protein sequences are diverse in their length of the N- and C-terminal regions and lack Q-rich stretches, with the exception of the *Chara braunii* (streptophyte algae) sequence featuring several extreme accumulations of Q.

The Arabidopsis LUG and LUH proteins have five WD domains in the C-terminal regions and a short N-terminal LisH domain required for SEU interaction. Additionally, LUG homologs feature a region, around 100 aa in length that is highly enriched in glutamine that is C-terminal of the LisH domain.. While all proteins in the phylogeny include two to five WD domains, the LisH domain is missing in a subset of proteins. All seed plants in our analysis include at least one LUG and one LUH representative with a similar protein structure, but may in some cases also encode proteins that lack parts present in the Arabidopsis sequences. Interestingly, *A. agrestis* LUG/LUH homolog lacks the LisH domain. However, among the streptophyte algae LUG/LUH homologs, only the Zygnematophyceae algae *Klebsormidium nitans* includes the LiSH domain, suggesting that LUG/SEU dimerization predates land plant origin and may have emerged in the Zygnematophyceae and was subsequently lost in the lineage leading to *A. agrestis.*

Protein interaction predictions show dimerization regions conserved across land plants

The structures of LUG and SEU protein family members have not been resolved yet and it remains unclear, which regions of the proteins physically interact. We were interested in the common principles of interaction of LUG and SEU proteins and carried out protein dimer predictions using Alphafold and identified regions mediating contact between LUG and SEU homologs (Fig. 3). While large portions of LUG and SEU homologs are composed of IDR that are unable to fold into 3D structures (Fig. 3A), the contact regions of LUG/SEU homolog dimers fold into alpha-helices (Fig. 3B-E). According to the structure predictions, LUG homologs form a long alpha-helix that is in contact with a shorter SEU homolog alpha-helix. In addition, at least three shorter alpha-helices of LUG homologs are in contact with three shorter alpha-helices of SEU homologs in the dimers consisting of members of the LUG and the SEU subfamily, and this is where the LisH domain is found in LUG homologs (Fig. 3B-D). In *P. patens* dimers, the shorter alpha-helices are not directly adjacent to the LisH domain (Fig. 3E). Interestingly, both *M. polymorpha* SEU homologs have a large surface composed of by beta-sheets in the region predicted to be in contact with MpLUG, which are not present in SEU homologs of vascular plants. The LisH domain of all LUG homologs is in contact with the SEU homologs in dimers, but in the tested proteins of *C. richardii* and the Bryophytes, additional contact is predicted with the WD region at the C-terminal part of the protein (Fig. 3F). All tested SEU homolog proteins interact with their respective LUG homolog partners via amino acid residues that are spread across more than 250 amino acids of the LIM-binding domain (Fig. 3F). This may require intricate protein folding of SEU homologs to enable proper positioning of these points of contact to LUG homologs.

LUG and SEU homologs are expressed uniformly in sporophytic and gametophytic tissue

Given the conserved protein domain structure and predicted protein interaction properties of LUG and SEU homologs, we assessed whether the functional diversification of LUG and SEU involves differential expression. We carried out digital gene expression analysis of *LUG* and *SEU* homologs in Arabidopsis, *C.richardii, P. patens,* and *M. polymorpha* of previously published transcriptomes. Fig. 4A shows that *LUH* of Arabidopsis is equally or stronger expressed than *LUG* in all tissues, while the *SEU* and *SLK1-3* genes are similar in expression strength, with *SLK2* being slightly stronger expressed. Both, *LUG* and SEU homologs are expressed in all tissues analyzed except for mature pollen.. Expression data of the histone modifiers *HDA19* and *SDG4* were added to the dataset.. *SDG4* is most strongly expressed stamen and mature pollen and floral organs, but expression is lacking from leaves and seeds. Interestingly, data from the ePlant browser (https://bar.utoronto.ca/eplant/; Waese et al. 2017) show expression of SGD4 not only in roots, but also in pistils, during stamen development and pollen tube growth through the ovary, suggesting that SGD4 may have an important role, maybe together with SEU, in gene activation in reproductive tissues, where they are both strongly expressed. In contrast, *HDA19* is expressed in all tissues analyzed, with least expression in pollen grains. The *C. richardii LUG* and and the *SEU* homologs are expressed uniformly throughout all analyzed tissues. *LUG* is substantially stronger expressed than the SEU homologs. Uniform expression through the analyzed tissues is also observed for the *P. patens LUG* and *SEU* genes (Fig. 4B). However, While *PpLUG1* is strongly expressed, *PpLUG2* expression in hardly detected. *PpLUG3, PpLUG4, PpSEU1* and *PpSEU2* are expressed at approximately the same levels with *PpSEU1* showing slightly higher expression and most genes show slightly higher expression in sporophytic than in gametophytic tissue (Fig. 4C). *MpLUG* of *M. polymorpha* is most strongly expressed in archegonia and shows least expression in sperm cells. *MpSEU1* and *MpSEU2* also show the weakest expression in archegonia.

Taken together, comparative expression analysis between land plant lineage representatives shows few differences in expression between LUG and SEU homologs, with the exception of PpLUG2, which is barely detectable. Taken together with the observed aberrant domain organization (Fig. 2) may be on its way to pseudogenization. This broad expression of LUG and SEU homologs across all land plant lineages is unexpected because many developmental regulators show expression restricted to specific tissues or developmental stages.

Acknowledgements: We thank all members of the FOR 5098 “ICIPS” for valuable discussion on the project. This work was supported by the de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and ELIXIR-DE (Forschungszentrum Jülich and W-de.NBI-001, W-de.NBI-004, W-de.NBI-008, W-de.NBI-010, W-de.NBI-013, W-de.NBI-014, W-de.NBI-016, W-de.NBI-022). This work was funded by the German Research Foundation (DFG), grant BE2547/26-1 to A. B.

Figure legends

Figure 1: Simplified phylogeny of SEU, SLK and pre-SEU/SLK proteins. Maximum likelihood phylogenetic tress were calculated with % bootstrap replicates of 1000 runs shown on the right side of the respective node. Values below 45 were collapsed. On the right side,

protein domain analyses are shown with amino acid number indicated on the Y axis. Q-residues, alpha-helices and the LIM-binding domain are shown.

Figures 2: Simplified phylogeny of LUG, LUH and pre-LUG/LUH proteins with % bootstrap replicates of 1000 runs shown on the right side of the respective node. Values below 45 were collapsed. On the right side, protein domain analyses are shown with amino acid number indicated on the Y axis. Q-residues, LisH and WD domains are shown.

Figure 3: LUG and SEU protein interaction prediction. Detailed overview over LUG (blue) and SEU (brown) amino acid residues with close proximity to interacting partner. **A)** - **E)** LUG (blue) and SEU (orange) protein complexes with close residues. Close residues are defined as residues with pIDDT confidence > 50 that are within 4.5 Å of a pIDDT > 50 confident partner chain residue. Regions with pIDDT < 50 are colored gray. For **B)** - **E)**, only the first 160 residues of LUG proteins are depicted, as well as SEU residues in proximity to LUG and 20 neighbouring atoms in both directions. **A)** AtLUG + AtSEU full-length protein dimer. Sites of contact of the AtLUG + AtSEU dimer **B)**; the CrLUG + CrSEU1 dimer **C)**; the MpLUG + MpSEU1 dimer **D)**; the PpLUG1 + PpSEU1 dimer **E)**. **F)** Position of contacting amino acid residues in the protein sequences in LUG (upper) and SEU (lower). Gray residues have either ≤ 50 pIDDT confidence, or have no close residue within 4.5 Å. Dark blue residues are within 4.5 Å to a residue of the other chain, with a pIDDT confidence of 50 < pIDDT ≤ 70 for both positions, while dark green residues have a pIDDT confidence of > 70.

Figure 4: Digital gene expression analysis of LUG and SEU homologs. Heatmaps of LUG and SEU expression values as log2(TPM + 1) in various developmental tissues. The scale is consistent across all subfigures. **A)** Expression values of adult *A. thaliana* leaves, roots and siliques, stage 10 embryos and stage 15 flowers and floral organs. Also containing expression values for SDG4 and HDAC19. **B)** *P. patens* expression values. **C)** Expression of orthologs in *C. richardii*. **D)** Expression in *M. polymorpha* Melbourne strain 15 d male and female thallus, Tak-1 14 d antheridiophores and antheridia, Tak-1 sperm cells, Tak-2 archegoniophores and archegonia, BC3 x Tak-1 13 d old sporophytes and in Cam1 x Cam2 spores.

References

Abramson, Josh; Adler, Jonas; Dunger, Jack; Evans, Richard; Green, Tim; Pritzel, Alexander et al. (2024): Accurate structure prediction of biomolecular interactions with AlphaFold 3. In: *Nature* 630 (8016), S. 493–500. DOI: 10.1038/s41586-024-07487-w.

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. (1990): Basic local alignment search tool. In: *Journal of molecular biology* 215 (3), S. 403–410. DOI: 10.1016/S0022-2836(05)80360-2.

Bao, Fang; Azhakanandam, Sridevi; Franks, Robert G. (2010): SEUSS and SEUSS-LIKE transcriptional adaptors regulate floral and embryonic development in Arabidopsis. In: *Plant physiology* 152 (2), S. 821–836. DOI: 10.1104/pp.109.146183.

Becker, Annette; Chen, Xia; Dresselhaus, Thomas; Gutsche, Nora; Müller-Schüssele, Stefanie J.; Sprunck, Stefanie et al. (2025): Sexual reproduction in land plants: an evolutionary perspective. In review at Plant Reproduction.

Bowman, John L.; Kohchi, Takayuki; Yamato, Katsuyuki T.; Jenkins, Jerry; Shu, Shengqiang; Ishizaki, Kimitsune et al. (2017): Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. In: *Cell* 171 (2), 287-304.e15. DOI: 10.1016/j.cell.2017.09.030.

Briginshaw, Liam N.; Flores-Sandoval, Eduardo; Dierschke, Tom; Alvarez, John P.; Bowman, John L. (2022): KANADI promotes thallus differentiation and FR-induced gametangiophore formation in the liverwort Marchantia. In: *The New phytologist* 234 (4), S. 1377–1393. DOI: 10.1111/nph.18046.

Conner, J.; Liu, Z. (2000): LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. In: *Proceedings of the National Academy of Sciences of the United States of America* 97 (23), S. 12902–12907. DOI: 10.1073/pnas.230352397.

Frank, Margaret H.; Scanlon, Michael 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. In: *Molecular biology and evolution* 32 (2), S. 355–367. DOI: 10.1093/molbev/msu303.

Franks, Robert G.; Wang, Chunxin; Levin, Joshua Z.; Liu, Zhongchi (2002): SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. In: *Development (Cambridge, England)* 129 (1), S. 253–263. DOI: 10.1242/dev.129.1.253.

Gonzalez, Deyarina; Bowen, Adam J.; Carroll, Thomas S.; Conlan, R. Steven (2007): The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. In: *Molecular and cellular biology* 27 (15), S. 5306–5315. DOI: 10.1128/MCB.01912-06.

Gregis, Veronica; Sessa, Alice; Colombo, Lucia; Kater, Martin M. (2006): AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. In: *The Plant cell* 18 (6), S. 1373–1382. DOI: 10.1105/tpc.106.041798.

Gregis, Veronica; Sessa, Alice; Dorca-Fornell, Carmen; Kater, Martin M. (2009): The Arabidopsis floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. In: *The Plant journal : for cell and molecular biology* 60 (4), S. 626–637. DOI: 10.1111/j.1365-313X.2009.03985.x.

Gu, Zuguang (2022): Complex heatmap visualization. In: *iMeta* 1 (3), e43. DOI: 10.1002/imt2.43.

Herrera-Ubaldo, Humberto; Campos, Sergio E.; López-Gómez, Pablo; Luna-García, Valentín; Zúñiga-Mayo, Víctor M.; Armas-Caballero, Gerardo E. et al. (2023): The protein-protein interaction landscape of transcription factors during gynoecium development in Arabidopsis. In: *Molecular plant* 16 (1), S. 260–278. DOI: 10.1016/j.molp.2022.09.004.

Higo, Asuka; Niwa, Masaki; Yamato, Katsuyuki T.; Yamada, Lixy; Sawada, Hitoshi; Sakamoto, Tomoaki et al. (2016): Transcriptional Framework of Male Gametogenesis in the Liverwort Marchantia polymorpha L. In: *Plant & cell physiology* 57 (2), S. 325–338. DOI: 10.1093/pcp/pcw005.

Hisanaga, Tetsuya; Fujimoto, Shota; Cui, Yihui; Sato, Katsutoshi; Sano, Ryosuke; Yamaoka, Shohei et al. (2021): Deep evolutionary origin of gamete-directed zygote activation by KNOX/BELL transcription factors in green plants. In: *eLife* 10. DOI: 10.7554/eLife.57090.

Julca, Irene; Ferrari, Camilla; Flores-Tornero, María; Proost, Sebastian; Lindner, Ann-Cathrin; Hackenberg, Dieter et al. (2021): Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. In: *Nature plants* 7 (8), S. 1143–1159. DOI: 10.1038/s41477-021-00958-2.

Katoh, Kazutaka; Misawa, Kazuharu; Kuma, Kei-ichi; Miyata, Takashi (2002): MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. In: *Nucleic acids research* 30 (14), S. 3059–3066. DOI: 10.1093/nar/gkf436.

Kolde, Raivo (2019): pheatmap: Pretty Heatmaps. Online verfügbar unter https://CRAN.R-project.org/package=pheatmap.

Lang, Daniel; Ullrich, Kristian K.; Murat, Florent; Fuchs, Jörg; Jenkins, Jerry; Haas, Fabian B. et al. (2018): The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. In: *The Plant journal : for cell and molecular biology* 93 (3), S. 515–533. DOI: 10.1111/tpj.13801.

Li, Fay-Wei; Nishiyama, Tomoaki; Waller, Manuel; Frangedakis, Eftychios; Keller, Jean; Li, Zheng et al. (2020): Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. In: *Nature plants* 6 (3), S. 259–272. DOI: 10.1038/s41477-020-0618-2.

Liu, Z.; Franks, R. G.; Klink, V. P. (2000): Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. In: *The Plant cell* 12 (10), S. 1879–1892. DOI: 10.1105/tpc.12.10.1879.

Liu, Z.; Meyerowitz, E. M. (1995): LEUNIG regulates AGAMOUS expression in Arabidopsis flowers. In: *Development (Cambridge, England)* 121 (4), S. 975–991. DOI: 10.1242/dev.121.4.975.

Liu, Zhongchi; Karmarkar, Vidyadhar (2008): Groucho/Tup1 family co-repressors in plant development. In: *Trends in plant science* 13 (3), S. 137–144. DOI: 10.1016/j.tplants.2007.12.005.

Marchant, D. Blaine; Chen, Guang; Cai, Shengguan; Chen, Fei; Schafran, Peter; Jenkins, Jerry et al. (2022): Dynamic genome evolution in a model fern. In: *Nature plants* 8 (9), S. 1038–1051. DOI: 10.1038/s41477-022-01226-7.

Mergner, Julia; Frejno, Martin; List, Markus; Papacek, Michael; Chen, Xia; Chaudhary, Ajeet et al. (2020): Mass-spectrometry-based draft of the Arabidopsis proteome. In: *Nature* 579 (7799), S. 409–414. DOI: 10.1038/s41586-020-2094-2.

Minh, Bui Quang; Schmidt, Heiko A.; Chernomor, Olga; Schrempf, Dominik; Woodhams, Michael D.; Haeseler, Arndt von; Lanfear, Robert (2020): Corrigendum to: IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. In: *Molecular biology and evolution* 37 (8), S. 2461. DOI: 10.1093/molbev/msaa131.

Parvathy, Jayadevan; Yazhini, Arangasamy; Srinivasan, Narayanaswamy; Sowdhamini, Ramanathan (2024): Interfacial residues in protein-protein complexes are in the eyes of the beholder. In: *Proteins* 92 (4), S. 509–528. DOI: 10.1002/prot.26628.

Pfannebecker, Kai C.; Lange, Matthias; Rupp, Oliver; Becker, Annette (2017): Seed Plant-Specific Gene Lineages Involved in Carpel Development. In: *Molecular biology and evolution* 34 (4), S. 925–942. DOI: 10.1093/molbev/msw297.

Pfluger, Jennifer; Zambryski, Patricia (2004): The role of SEUSS in auxin response and floral organ patterning. In: *Development (Cambridge, England)* 131 (19), S. 4697–4707. DOI: 10.1242/dev.01306.

Posit team (2024): RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC. Online verfügbar unter http://www.posit.co/.

R Core Team (2022): R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Online verfügbar unter https://www.R-project.org/.

Roessner, Clemens; Griep, Sven; Becker, Annette (2024): A land plant phylogenetic framework for GLABROUS INFLORESCENCE STEMS (GIS), SUPERMAN, JAGGED and allies plus their TOPLESS co-

repressor. In: *Molecular phylogenetics and evolution* 201, S. 108195. DOI: 10.1016/j.ympev.2024.108195.

Schrödinger LLC (2010): The PyMOL Molecular Graphics System, Version 3.0.

Shi, Lei; Li, Changjiang; Lv, Gaofeng; Li, Xing; Feng, Wutao; Bi, Yujing et al. (2024): The adaptor protein ECAP, the corepressor LEUNIG, and the transcription factor BEH3 interact and regulate microsporocyte generation in Arabidopsis. In: *The Plant cell* 36 (7), S. 2531–2549. DOI: 10.1093/plcell/koae086.

Simon, Marissa K.; Skinner, Debra J.; Gallagher, Thomas L.; Gasser, Charles S. (2017): Integument Development in Arabidopsis Depends on Interaction of YABBY Protein INNER NO OUTER with Coactivators and Corepressors. In: *Genetics* 207 (4), S. 1489–1500. DOI: 10.1534/genetics.117.300140.

Sitaraman, Jayashree; Bui, Minh; Liu, Zhongchi (2008): LEUNIG_HOMOLOG and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. In: *Plant physiology* 147 (2), S. 672–681. DOI: 10.1104/pp.108.115923.

Sreedasyam, Avinash; Plott, Christopher; Hossain, Md Shakhawat; Lovell, John T.; Grimwood, Jane; Jenkins, Jerry W. et al. (2023): JGI Plant Gene Atlas: an updateable transcriptome resource to improve functional gene descriptions across the plant kingdom. In: *Nucleic acids research* 51 (16), S. 8383–8401. DOI: 10.1093/nar/gkad616.

Sridhar, Vaniyambadi V.; Surendrarao, Anandkumar; Gonzalez, Deyarina; Conlan, R. Steven; Liu, Zhongchi (2004): Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. In: *Proceedings of the National Academy of Sciences of the United States of America* 101 (31), S. 11494–11499. DOI: 10.1073/pnas.0403055101.

Sridhar, Vaniyambadi V.; Surendrarao, Anandkumar; Liu, Zhongchi (2006): APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. In: *Development (Cambridge, England)* 133 (16), S. 3159–3166. DOI: 10.1242/dev.02498.

Su, Zhao; Ma, Xuan; Guo, Huihong; Sukiran, Noor Liyana; Guo, Bin; Assmann, Sarah M.; Ma, Hong (2013): Flower development under drought stress: morphological and transcriptomic analyses reveal acute responses and long-term acclimation in Arabidopsis. In: *The Plant cell* 25 (10), S. 3785–3807. DOI: 10.1105/tpc.113.115428.

Waese, Jamie; Fan, Jim; Pasha, Asher; Yu, Hans; Fucile, Geoffrey; Shi, Ruian et al. (2017): ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. In: *The Plant cell* 29 (8), S. 1806–1821. DOI: 10.1105/tpc.17.00073.

Wang, Boyu; Zhang, Honghong; Huai, Junling; Peng, Fangyu; Wu, Jie; Lin, Rongcheng; Fang, Xiaofeng (2022): Condensation of SEUSS promotes hyperosmotic stress tolerance in Arabidopsis. In: *Nature chemical biology* 18 (12), S. 1361–1369. DOI: 10.1038/s41589-022-01196-z.

Wang, Zhen; Lu, Kaiheng; Liu, Xuan; Zhu, Yuping; Liu, Changhong (2023): Comparative Functional Genome Analysis Reveals the Habitat Adaptation and Biocontrol Characteristics of Plant Growth-Promoting Bacteria in NCBI Databases. In: *Microbiology spectrum* 11 (3), e0500722. DOI: 10.1128/spectrum.05007-22.

Wickham, Hadley (2016): ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York. Online verfügbar unter https://ggplot2.tidyverse.org.

Zhai, Huawei; Zhang, Xiaoyue; You, Yanrong; Lin, Lihao; Zhou, Wenkun; Li, Chuanyou (2020): SEUSS integrates transcriptional and epigenetic control of root stem cell organizer specification. In: *The EMBO journal* 39 (20), e105047. DOI: 10.15252/embj.2020105047.

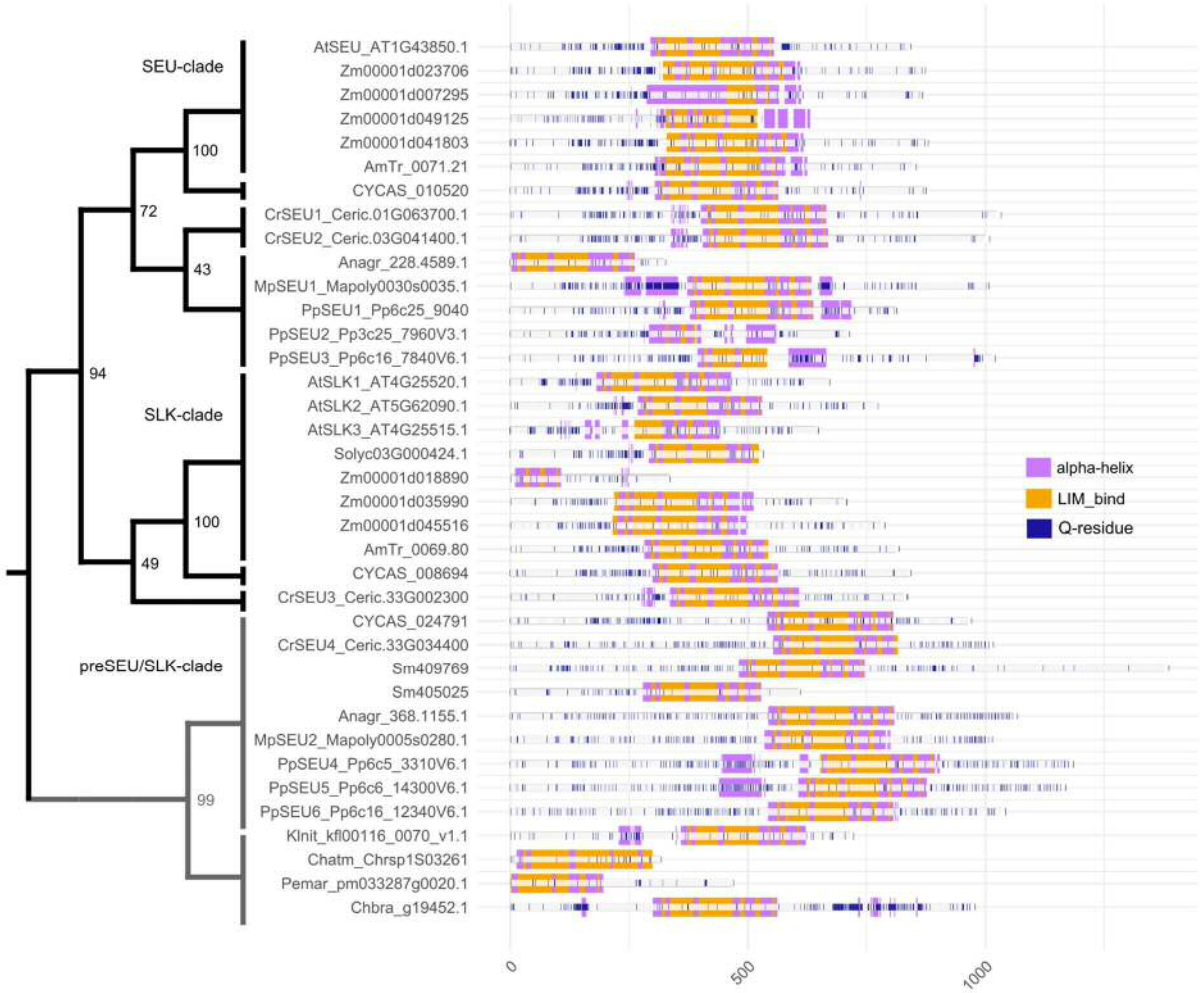
SEU-clade
SLK-clade
preSEU/SLK-clade
100
72
43
94
100
49
99
AtSEU_AT1G43850.1
Zm00001d023706
Zm00001d007295
Zm00001d049125
Zm00001d041803
AmTr_0071.21
CYCAS_010520
CrSEU1_Ceric.01G063700.1
CrSEU2_Ceric.03G041400.1
Anagr_228.4589.1
MpSEU1_Mapoly0030s0035.1
PpSEU1_Pp6c25_9040
PpSEU2_Pp3c25_7960V3.1
PpSEU3_Pp6c16_7840V6.1
AtSLK1_AT4G25520.1
AtSLK2_AT5G62090.1
AtSLK3_AT4G25515.1
Solyc03G000424.1
Zm00001d018890
Zm00001d035990
Zm00001d045516
AmTr_0069.80
CYCAS_008694
CrSEU3_Ceric.33G002300
CYCAS_024791
CrSEU4_Ceric.33G034400
Sm409769
Sm405025
Anagr_368.1155.1
MpSEU2_Mapoly0005s0280.1
PpSEU4_Pp6c5_3310V6.1
PpSEU5_Pp6c6_14300V6.1
PpSEU6_Pp6c16_12340V6.1
Klnit_kfl00116_0070_v1.1
Chatm_Chrsp1S03261
Pemar_pm033287g0020.1
Chbra_g19452.1
alpha-helix
LIM_bind
Q-residue
0
500
1000

Figure 1

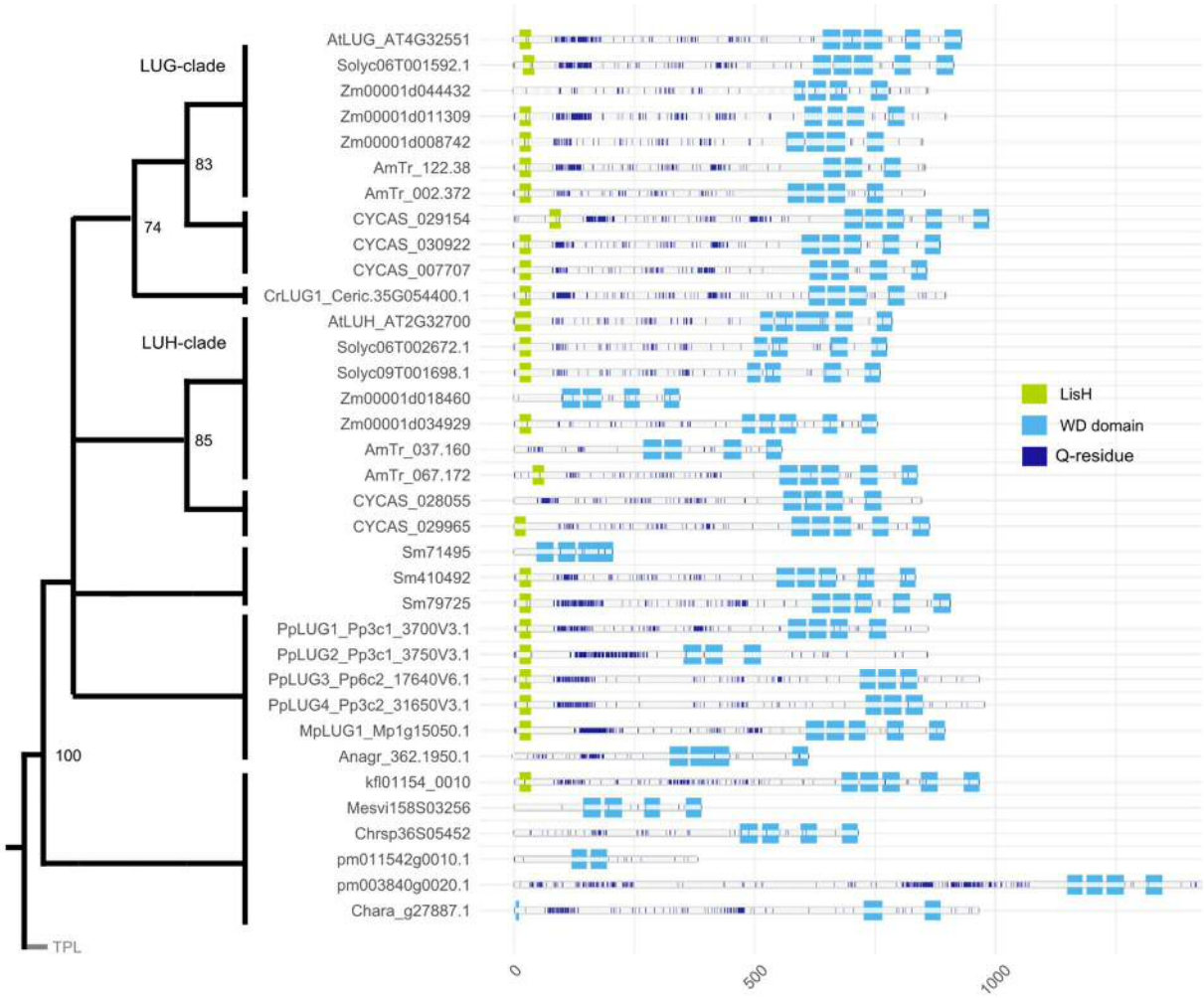
LUG-clade
LUH-clade
83
74
85
100
TPL
AtLUG_AT4G32551
Solyc06T001592.1
Zm00001d044432
Zm00001d011309
Zm00001d008742
AmTr_122.38
AmTr_002.372
CYCAS_029154
CYCAS_030922
CYCAS_007707
CrLUG1_Ceric.35G054400.1
AtLUH_AT2G32700
Solyc06T002672.1
Solyc09T001698.1
Zm00001d018460
Zm00001d034929
AmTr_037.160
AmTr_067.172
CYCAS_028055
CYCAS_029965
Sm71495
Sm410492
Sm79725
PpLUG1_Pp3c1_3700V3.1
PpLUG2_Pp3c1_3750V3.1
PpLUG3_Pp6c2_17640V6.1
PpLUG4_Pp3c2_31650V3.1
MpLUG1_Mp1g15050.1
Anagr_362.1950.1
kfl01154_0010
Mesvi158S03256
Chrsp36S05452
pm011542g0010.1
pm003840g0020.1
Chara_g27887.1
0
500
1000
LisH
WD domain
Q-residue

Figure 2

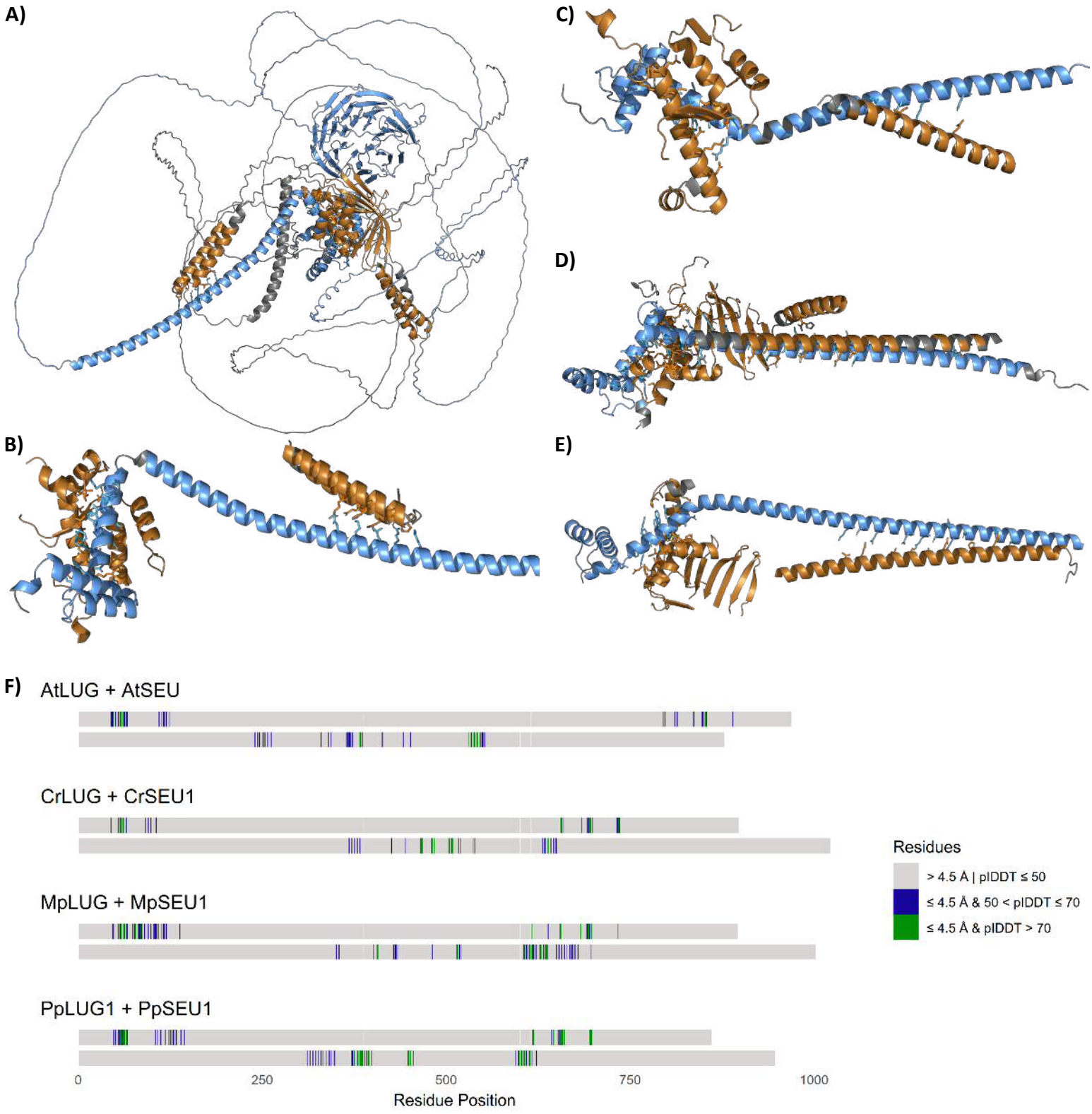
A)
C)
D)
B)
E)
F)
AtLUG + AtSEU
CrLUG + CrSEU1
MpLUG + MpSEU1
PpLUG1 + PpSEU1
0
250
500
750
1000
Residue Position
Residues
> 4.5 Å | plDDT ≤ 50
≤ 4.5 Å & 50 < plDDT ≤ 70
≤ 4.5 Å & plDDT > 70

Figure 3

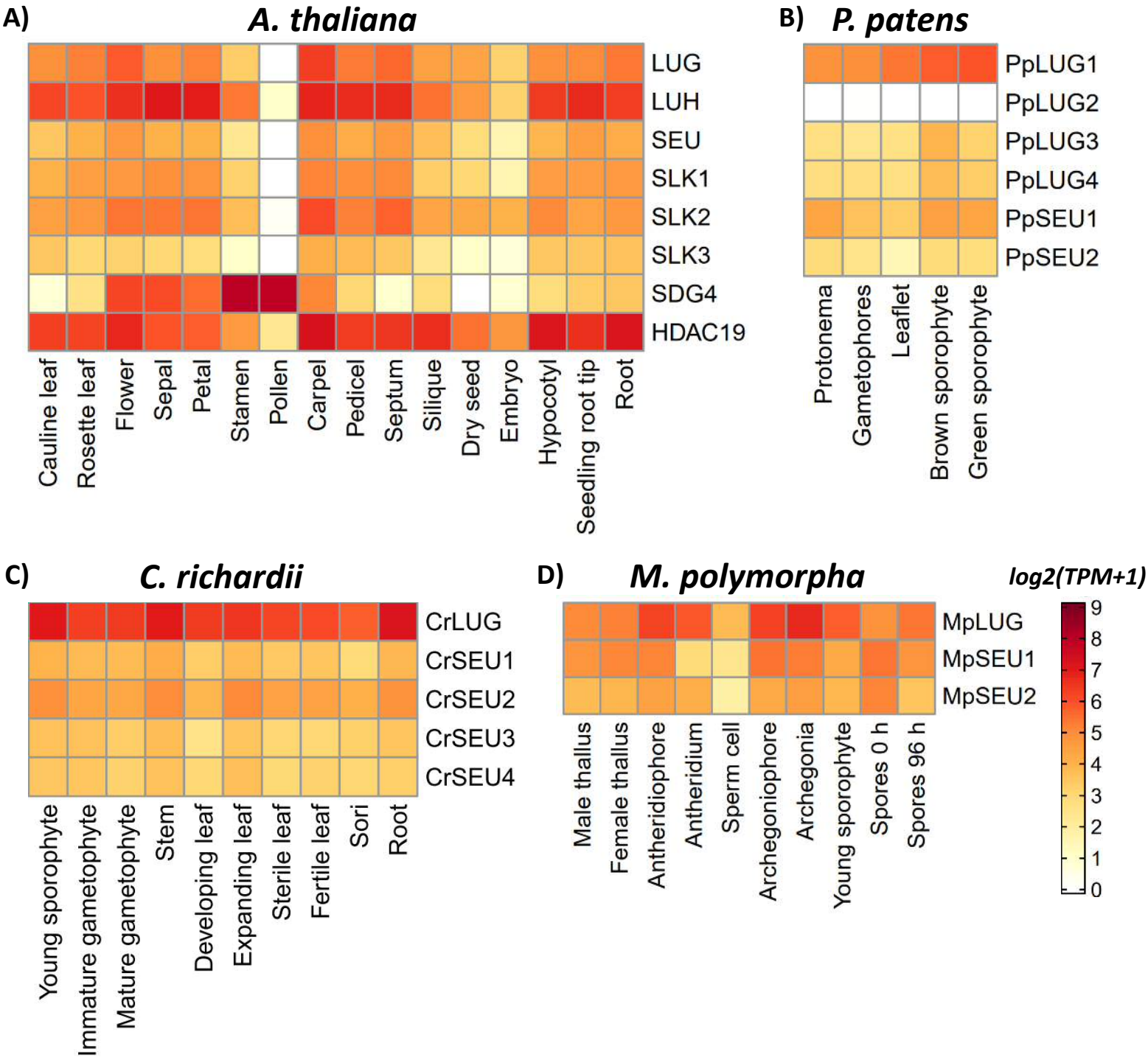
A)
A. thaliana
LUG
LUH
SEU
SLK1
SLK2
SLK3
SDG4
HDAC19
Cauline leaf
Rosette leaf
Flower
Sepal
Petal
Stamen
Pollen
Carpel
Pedicel
Septum
Silique
Dry seed
Embryo
Hypocotyl
Seedling root tip
Root
B)
P. patens
PpLUG1
PpLUG2
PpLUG3
PpLUG4
PpSEU1
PpSEU2
Protonema
Gametophores
Leaflet
Brown sporophyte
Green sporophyte
C)
C. richardii
CrLUG
CrSEU1
CrSEU2
CrSEU3
CrSEU4
Young sporophyte
Immature gametophyte
Mature gametophyte
Stem
Developing leaf
Expanding leaf
Sterile leaf
Fertile leaf
Sori
Root
D)
M. polymorpha
MpLUG
MpSEU1
MpSEU2
Male thallus
Female thallus
Antheridiophore
Antheridium
Sperm cell
Archegoniophore
Archegonia
Young sporophyte
Spores 0 h
Spores 96 h
log2(TPM+1)
9
8
7
6
5
4
3
2
1
0

Figure 4