Systematic Analysis of the Functional Overlap among Plant YTHDF proteins

Daniel Flores-Téllez, Mathias Tankmar, Sören von Bülow, Junyu Chen, Kresten Lindorff-Larsen, Peter Brodersen* and Laura Arribas-Hernández*

University of Copenhagen, Copenhagen Plant Science Center, Ole Maaløes Vej 5, DK-2200 Copenhagen N

*Corresponding authors

Email pbrodersen@bio.ku.dk (P.B.) and laura.arribas@bio.ku.dk (L.A.-H.)

Abstract

YT521-B homology (YTH) domain proteins act as readers of *N6*-methyladenosine (m^6A), the most common internal covalent modification in eukaryotic mRNA. Members of the YTHDF subclade can determine properties of m^6A-containing mRNAs in the cytoplasm. Vertebrates encode three YTHDF proteins whose possible functional specialization is debated. In land plants, the YTHDF clade has expanded from one member in basal lineages to eleven so-called EVOLUTIONARILY CONSERVED C-TERMINAL REGION1-11 (ECT1-11) proteins in *Arabidopsis thaliana*, named after the conserved YTH domain placed behind a long N-terminal intrinsically disordered region (IDR). *ECT2*, *ECT3* and *ECT4* show genetic redundancy in stimulation of primed stem cell division, but the origin and implications of YTHDF expansion in higher plants as a whole are unknown, as it is unclear whether it involves acquisition of fundamentally different properties, in particular of their divergent IDRs. Here, we used functional complementation of *ect2/ect3/ect4* mutants to test whether different YTHDF proteins can perform the same function if expressed at similar levels in leaf primordia. We show that stimulation of primordial cell division relies on an ancestral molecular function of the m^6A-YTHDF axis in land plants that is present in bryophytes and is conserved over YTHDF diversification, as it appears in all major clades of YTHDF proteins in flowering plants. Nevertheless, lineage-specific neo-functionalization of ECT1, ECT9 and ECT11 happened after late duplication events, and involves altered properties especially of the IDRs, but also of the YTH domains. Metazoan YTHDFs cannot replace ECT2/3/4 function in arabidopsis, perhaps suggesting different molecular activities of those YTHDF proteins.

Significance statement

Regulation of gene expression is essential to life. It ensures correct balancing of cellular activities and the controlled proliferation and differentiation necessary for the development of multicellular organisms. Among other mechanisms, gene expression can be controlled by methylation of adenosines in mRNA (m^6A). Absence of m^6A impairs embryo development in plants and vertebrates, and its mis-regulation is associated with human cancers. m^6A-dependent regulation can be exerted by a group of cytoplasmic proteins called YTHDFs. Plants have many more YTHDFs than animals, but the reason for this is unknown. This study addresses the origin of plant YTHDF expansion, and reveals that most, but not all, have the same molecular functions that facilitate rapid division of differentiating stem cells.

Introduction

N6-methyladenosine (m^6A) is the most abundant modified nucleotide occurring internally in eukaryotic mRNA. It is of major importance in gene regulation as illustrated by the embryonic lethality of mutants in the dedicated mRNA adenosine methyltransferase in higher plants (1) and in mammals (2). The presence of m^6A in an mRNA may have multiple biochemical consequences. These include changes in the secondary structure (3-5) and the creation of binding sites for RNA-binding proteins specialized for m^6A recognition (6, 7). YT521-B homology (YTH) domain proteins (8) constitute the best studied class of m^6A-binding proteins (6, 7, 9). They achieve specificity for m^6A via an aromatic pocket accommodating the *N6*-adenosine methyl group, such that the affinity of isolated YTH domains for m^6A-containing RNA is 10-20-fold higher than for unmodified RNA (9-14).

Two different phylogenetic clades of YTH domains have been defined, YTHDC and YTHDF (6, 15). Genetic studies establish that major functions of m^6A in development in both vertebrates and plants depend on the YTHDF clade of readers (16-22). In all cases studied in detail thus far, plant and animal YTHDF proteins are cytoplasmic in unchallenged conditions (15, 16, 23-26), and contain a long N-terminal intrinsically disordered region (IDR) in addition to the C-terminal YTH domain (6, 8, 27). While the YTH domain is necessary for specific binding to m^6A in mRNA (9-14), the IDR is considered to be the effector part of the protein (9, 28-30). Nonetheless, it has been proposed that the IDR may also participate in RNA binding, because the YTH domain alone has low affinity for mRNA (6). Indeed, the IDR-dependent crosslinks between a YTHDF protein and mRNA detected upon UV-irradiation of living arabidopsis seedlings (31) experimentally supports such a mechanism, conceptually equivalent to the contribution of IDRs in transcription factors to specific DNA binding (32).

While yeast, flies and primitive land plants encode only one YTHDF protein (24-26, 33, 34), vertebrates have three closely related paralogs (YTHDF1-3) (6, 8, 15, 35) and higher plants encode an expanded family, with eleven members in *Arabidopsis thaliana* (*Ath*) referred to as EVOLUTIONARILY CONSERVED C-TERMINAL REGION1-11 (ECT1-11) (27, 36). It is a question of fundamental importance for the understanding of how complex eukaryotic systems use the regulatory potential of m^6A whether these many YTHDF proteins perform the same biochemical function, or whether their molecular properties are specialized. Such molecular specialization could, for instance, arise as a consequence of differential binding specificity to mRNA targets, distinct protein-protein interaction properties, or distinct biophysical properties of IDRs, such as those related to the propensity to phase separate (37, 38). Alternatively, diversification of biological functions could be achieved with distinct expression patterns or induction by environmental cues, even if targets and molecular functions are the same.

Initial studies on mammalian cell cultures advocated a model in which YTHDF1 would enhance translation of target mRNAs, YTHDF2 would promote mRNA decay, and YTHDF3 would be able to trigger either of the two (9, 39-41). Nonetheless, recent studies in mouse, zebrafish and human cell culture involving single and combined *ythdf* knockouts and analysis of interacting mRNAs and proteins do not support functional specialization, and propose a unified molecular function for all three vertebrate YTHDFs in accelerating mRNA decay (18, 22, 35).

The great expansion of YTHDF proteins in higher plants is unique among eukaryotes (8). Phylogenetic analyses of plant YTHDF domains have established the existence of 3 clades in angiosperms, DF-A (comprising *Ath* ECT1, ECT2, ECT3, ECT4), DF-B (comprising *Ath* ECT5, ECT10, ECT9), and DF-C (comprising *Ath* ECT6, ECT7, ECT8 and ECT11) (26). The fact that the eleven arabidopsis paralogs have the aromatic residues necessary for specific binding to m^6A (42) suggests that they may all function as m^6A readers. The unified model for YTHDF function recently proposed for the three vertebrate paralogs (18, 22, 35) is consistent with what had already been established for *Ath* ECT2, ECT3 and ECT4 in the plant DF-A clade (16), even though the three plant paralogs are more divergent in sequence than the highly similar mammalian YTHDF1-3 (8, 16, 35). The two most highly expressed members in *A. thaliana*, ECT2 and ECT3, accumulate in dividing cells of organ primordia and exhibit genetic redundancy in the stimulation of stem cell proliferation during organogenesis (16, 17). The two proteins probably act truly redundantly *in vivo* to control this process, because they associate with highly overlapping target sets in wild type plants, and each exhibits increased target mRNA occupancy in the absence of the other protein (23). Simultaneous knockout of *ECT2* and *ECT3* causes a 2-day delay in the emergence of the first true leaves, aberrant leaf morphology, slow root growth and defective root growth directionality among other defects (16, 17) that resemble those of plants with diminished m^6A deposition (43-45). For the third DF-A clade member, *ECT4*, the genetic redundancy is only noticeable in some tissues as an exacerbation of *ect2/ect3* phenotypes upon additional mutation of *ect4*, most conspicuously seen in leaf morphogenesis (16, 17). Despite the strong evidence for redundant functions among these three YTHDF paralogs in the DF-A clade, the presence of many other YTHDF proteins in arabidopsis leaves open the question of whether substantial functional specialization of YTHDF proteins exists in plants.

In this study, we systematically define overlaps in molecular functions of plant YTHDF proteins. Employing functional complementation of leaf formation in triple *ect2/ect3/ect4* (*te234*) mutants, we demonstrate that at least one member of all clades in arabidopsis, and the only YTHDF protein from a group of primitive land plants, are able to replace ECT2/3/4 functions. In contrast, a few late-diverging ECTs were not able to perform the molecular functions of ECT2/3/4. Based on these results, we propose an ancestral

molecular role for land plant YTHDF proteins in stimulation of primordial cell proliferation, and sustained functional redundancy during the diversification process that started more than 400 million years ago (Mya). In addition, our results also support the functional divergence of a small subset of fast-evolving plant YTHDF proteins with contributions to specialization mainly from the IDRs, but also from the YTH domains.

Results

The phylogeny of plant YTHDF proteins comprises several clades

The adaptation of plants to terrestrial life, since the transition from freshwater to land until the rise of angiosperms, was accompanied by the acquisition of morphological, physiological and genomic complexity (46, 47). Knowing how diversification of YTHDF proteins came about during the course of plant evolution is relevant to understand their distinct functions, and may hint at roles they might have played in this process. However, the species included in the so far most detailed phylogenetic study on plant YTHDF proteins jump from bryophytes (mosses, liverworts and hornworts (47-52)) to angiosperms (flowering plants) (26). To include the YTHDF repertoires of intermediately diverging clades such as ferns and gymnosperms, we performed a phylogenetic analysis using YTHDF proteins from taxa widely spread across terrestrial plant evolution (Fig. 1*A*). When possible, we included basal clades of each group in addition to model organisms. The resulting phylogenetic tree largely agrees with the DF-A/B/C clades defined on the basis of angiosperm YTHDF sequences (26), with some minor differences. We consider that the former DF-C group comprising *Ath* ECT6/7/8/11 (26) can be subdivided into two groups that diverged early during YTHDF radiation (approximate likelihood-ratio test (aLRT) = 87), because the split between them is comparable to that of clades DF-A and -B (Fig. 1*B*). Thus, we introduced the name 'DF-D' for the clade defined by one *Amborella trichopoda* (a basal angiosperm) YTHDF protein (*Atr* DFD) and *Ath* ECT6/7 (Fig. 1*B*). Furthermore, since the group comprising bryophyte YTHDFs did not receive a designation in previous studies, we named it DF-E (mnemotechnic for 'Early') (Fig. 1*B*). Finally, an additional group composed of early-diverging YTHDF proteins from the fern *Ceratopteris richardii* (*Cri* DF1/2/3) and the gymnosperm *Ginkgo biloba* (*Gb* DF1/2) was named DF-F, as it did not fall into any of the other clades (Fig. 1*B*). We conclude that land plant YTHDFs are phylogenetically more diverse than previously appreciated.

YTHDF protein diversification occurred early during land plant evolution

Our phylogenetic analysis shows that plant YTHDF diversification started at least before the radiation of Euphyllophytes (plants with true leaves comprising ferns, gymnosperms and

angiosperms) more than 410 Mya (51, 53). This is because bryophytes possess one or two YTHDFs in the 'early clade' (DF-E) while the fern *C. richardii* has six YTHDFs in 3 different clades (Fig. 1*B*). Thus, our analysis reveals that YTHDF radiation started early in land plant evolution and coincided with the acquisition of morphological complexity and the adaptation to diverse environments.

A functional complementation assay for plant YTHDF proteins

To address the degree of functional specialization among plant YTHDF proteins in a simple manner, we set up a functional complementation assay aimed to determine which of the eleven arabidopsis YTHDF proteins (Fig. 2*A*) are able to perform the molecular functions of ECT2/3/4 required for rapid cellular proliferation in leaf primordia. Initial attempts at using the *ECT2* promoter for ectopic *ECT* expression in primordial cells were unsuccessful, as pilot experiments with a genomic *ECT4* fragment revealed that its expression was substantially lower than that of a similar *ECT2* genomic fragment when driven by the *ECT2* promoter (Fig. S1), perhaps pointing to the presence of internal *cis*-regulatory elements. We therefore turned to the use of *ECT* cDNAs (c*ECT*s) under the control of the promoter of the ribosomal protein-encoding gene *AthuS7B*/*RPS5A* (At3g11940 (54)) active in dividing cells (55) in a pattern that resembles the *ECT2/3/4* expression domain (16, 17). Transformation of the *uS7Bp:cECT2-mCherry* construct in *te234* plants resulted in complementation frequencies of ~40-50% among primary transformants, a percentage slightly lower but comparable to that obtained with the genomic *ECT2-mCherry* construct under the control of *ECT2* promoter and terminator regions (*ECT2p:gECT2-mCherry*) (16) (Fig. S2*A*). Free mCherry expressed from a *uS7Bp:mCherry* transgene showed no complementation (Fig. S2*A*), as expected. Importantly, the leaf morphology and the pattern of mCherry fluorescence in *uS7Bp:cECT2-mCherry* lines was indistinguishable from that of *ECT2p:gECT2-mCherry* (Fig. S2*B*). Hence, we proceeded with expression of cDNAs encoding all of ECT1-11 fused to C-terminal mCherry under the control of the *uS7B* promoter (henceforth ECT1-11) in *te234* mutants (Fig. 2*B*).

Percentage of complementation among primary transformants as a readout for functionality

To design the experimental approach in detail, we also considered the fact that expression of transgenes involves severe variations in levels and patterns among transformants. This is due to positional effects of the T-DNA insertion and the propensity of transgenes to trigger silencing in plants (56), and explains why only a fraction of *ECT2-mCherry* lines are able to rescue loss of ECT2 function (Fig. S2*A*, (16)). Hence, many independent lines need to be analyzed before choosing stable and representative lines for further studies. Taking this into account, we decided to perform systematic and unbiased comparisons on ECT functionality

by counting the fraction of primary transformants (henceforth T1s) able to complement the late leaf emergence of *te234* plants for each construct. We set a size threshold for the rapid assessment of large numbers of transformants considering that leaves longer than 0.5 mm after 10 days of growth constitute unambiguous sign of complementation, because this is the minimum size never reached by *te234* or control lines (Fig. S2*B*). Hence, we calculated complementation percentages as the number of T1s with leaf size (s) larger than 0.5 mm relative to the total number of transformants, and averaged that score over several independent transformations (Figs. 2*C*, S3*A*, and S4*A*).

Most arabidopsis YTHDF paralogs can perform the molecular function required for leaf development

The results of the comparative analysis reveal a high degree of functional overlap within the arabidopsis YTHDF family, because eight out of the eleven YTHDF proteins, namely ECT2/3/4/5/6/7/8/10, are able to complement delayed leaf emergence of *ect2/3/4* mutants with frequencies ranging from 14% (ECT6) to 39% (ECT2) when expressed in the same cells (Figs. 2*C* and S4*A*). Importantly, complementation is indirect proof of m^6A-binding ability, because the proteins must bind to the targets whose m^6A-dependent regulation is necessary for correct leaf organogenesis in order to restore ECT2/3/4 function. That is not trivial, because some eukaryotes lacking canonical m^6A-methyltransferases encode YTH domain proteins (57). Indeed, the YTH domain of fission yeast Mmi1 does not bind to m^6A (58). Thus, most, but not all, arabidopsis YTHDF proteins retain the molecular functions required to stimulate proliferation of primed stem cells. Of the three remaining proteins, ECT1 showed only residual complementation capacity, while no complementation was observed for ECT11 at this stage (Figs. 2*C* and S4*A*). Finally, we could not conclude at this point whether ECT9 could replace ECT2/3/4 function or not, because this transgene produced very few transformants in the *te234* background (Figs. 2*C*, S3*A* and S5) with barely visible fluorescence. All such transformants underwent complete silencing in the following generation (T2).

Differences in complementation scores are not explained by a bias in expression levels

Our T1 scoring method relies on the assumption that the *uS7B* promoter drives similar ranges of expression of the different *ECT* cDNAs. However, different silencing propensities between the *ECT* cDNAs might introduce bias in the results. To take this possibility into account, we first performed western blotting from three lines per transgene selected as the best-complementing in their group (Fig. S4*B*). The results showed that the higher complementation scores of e.g. ECT2, ECT3 and ECT8 are not explained by higher expression. Rather, the reverse was true. For example, while ECT2 could complement *te234*

plants with the lowest expression among all ECTs, the few lines expressing ECT1 able to pass the complementation threshold had the highest expression of all ECTs, suggesting that ECT1 can only partially compensate for loss of ECT2/3/4 when overexpressed.

The same functional classes of ECTs emerge from complementation of silencing-deficient te234

Next, we tested whether the same results could be reproduced in plants lacking *RNA-DEPENDENT RNA POLYMERASE 6* (*RDR6*), a gene essential for the biogenesis of transgene-derived short interfering RNAs (siRNAs) (59, 60). Because siRNA-guided transgene silencing is impaired in *rdr6* mutants, higher and more homogenous ranges of expression among different transgenes can be expected in this background. Hence, we constructed *rdr6-12/te234* mutants that exhibited both the narrow, downward-curling leaf phenotype of *rdr6* mutants (61) and the characteristic rosette morphology of *te234* (16) (Fig. S6*A*). Because the delay in leaf emergence was identical in *te234* and *rdr6-12/te234* mutants (Fig. S6*A*) and the complementation rates of *ECT2* transgenes were much higher in the silencing-deficient plants (Fig. S2*A*), we proceeded to repeat the functional assay in the *rdr6-12/te234* background. The results showed an overwhelming increase in the complementation score of the eight functionally equivalent ECTs (ECT2/3/4/5/6/7/8/10), now ranging from 59% (ECT6) to 94% (ECT2). The relative differences between these ECTs were less pronounced, but followed a trend largely unchanged compared to the *te234* background (Fig. 2*C,D*). On the other hand, the difference between the eight ECTs with ECT2/3/4-like function and the group of ECTs with negligible complementation capacity (ECT1/9/11) was highlighted by the *rdr6* mutation: ECT1 and ECT11 did not increase their scores (Fig. 2*D*), and absence of silencing resulted in a number of *ECT9* transformants with visible mCherry fluorescence, yet no complementation activity (Figs. 2*D* and S6*B*), thus verifying the lack of ECT2/3/4-like function of ECT9.

Rosette phenotypes corroborate the complementation scores

To verify the results of the complementation assay in the adult phenotype, we characterized rosettes of two independent lines for each transgene (Fig. 2*E-H*). As expected, *rdr6-12/te234* plants expressing ECTs with high complementation scores resembled single *rdr6-12* mutants (Fig. 2*E,F*), while the ECT1 and ECT11 transformants with the biggest first true leaves at 10 DAG produced plants only slightly bigger than the *rdr6-12/te234* background (Fig. 2*E,G*), indicating low but residual ECT2/3/4-like function for these two proteins. Strikingly, *rdr6-12/te234* plants expressing ECT9 exhibited strong developmental phenotypes (Figs. 2*H* and S6*B*) that resembled those of methyltransferase deficient plants such as *amiR-MTA* (44) or *vir-1* (45). Such an effect was not observed for any other *Ath*

ECT. This result suggests that ECT9 can bind to m^6A targets, either exerting opposite molecular actions compared to ECT2/3/4, or simply precluding their possible regulation by remaining ECTs such as ECT5/6/7/8/10 by competitive binding. We tested whether ECT9 can outcompete endogenous ECT2/3/4 in the proliferating cells of the apex by expressing the ECT9 transgene in single *rdr6-12* mutant plants. We found no evidence for this, because the transformation produced a high amount of fluorescent T1s without obvious defects (Fig. S7).

Altogether, our results indicate that arabidopsis YTHDF proteins may be divided into three functional classes, i) ECT2/3/4/5/6/7/8/10 with the molecular functionality of ECT2/3/4 required for correct and timely leaf formation; ii) ECT1/11, with only residual ECT2/3/4 functions, and iii) ECT9, unable to perform such functions and toxic when expressed in the expression domain of ECT2/3/4 in their absence.

Poor ECT2/3/4-like activity of ECT1 and ECT11 IDRs is the main reason for their reduced or absent complementation capacity of te234 plants

To investigate whether lack of *te234* mutant complementation by ECT1/9/11 is due to poor or divergent target-binding affinity by the YTH domain, IDR-related effector functions, or a combination of both, we built chimeric IDR/YTH constructs between ECT2 and ECT1/9/11 (Fig. 3*A*). Additionally, we built an ECT8$_{IDR}$/ECT2$_{YTH}$ chimera (C-8/2) to test whether hybrid constructs can at all be functional, choosing ECT8 as a positive control due to its high scores in our functional assay (Fig. 2*C,D*) and the robust morphological complementation of adult plants (Fig. 2*F*). Expression of the constructs in *te234* mutants showed that C-8/2 chimeras were fully functional (Figs. 3*B, S3B* and *S8*), thus validating the approach. Expression of ECT1- and ECT11-derived chimeras showed that both the IDR and the YTH domain of the two proteins retain a reduced degree of function compared to the equivalent regions in ECT2 (Figs. 3*B, S3B* and *S8*). However, the IDRs of both proteins scored lower than their YTH domains when combined with the other half from ECT2 (Fig. 3*B*), pointing to the IDR as the main cause for different functionality. In the wild type ECT1 and ECT11 proteins, the combination of the two poorly-performing parts is likely the cause for the minute ECT2/3/4-like activity of ECT1 and ECT11.

The IDR of ECT9 is incapable of performing ECT2/3/4 molecular functions, but its YTH domain retains some ECT2/3/4-like function

Transformation with the ECT2$_{IDR}$/ECT9$_{YTH}$ chimera (C-2/9) resulted in a normal number of T1s of which up to 17% showed signs of complementation compared to 56% for ECT2, suggesting that the YTH domain of ECT9 also retains at least partial functionality. However, the reverse construct (ECT9$_{IDR}$/ECT2$_{YTH}$, C-9/2)) recapitulated the results obtained for ECT9:

low transformation efficiency and low fluorescence in the few recovered T1s, that underwent complete silencing in the next generation. Hence, we introduced the C-9/2 chimera into the *rdr6-12/te234* background and observed plants with stronger fluorescence exhibiting aberrant morphology (Figs. 3*C,D*, S3*C* and S6*B*), similar to what was obtained with ECT9. These results confirm that ECT9 is not functionally equivalent to ECT2/3/4, and point to the IDR as the main site of functional divergence of the protein.

The molecular requirements for m^{6}A-binding are conserved in the YTH domains of ECT1/9/11

To pinpoint molecular characteristics of ECT1/9/11 responsible for their different functionality, we performed a series of comparative analyses between ECT1/9/11 and other ECT proteins. We first analyzed the YTH domains, taking ECT2 as a reference and using the structurally well-characterized YTH domain of *Homo sapiens* (*Hs*) YTHDF1 in complex with RNA GGm^{6}ACU (62) as a model. Amino acid sequence alignment of the YTH domains of all arabidopsis ECTs (Fig. S9), AlphaFold-predictions (63, 64) of the apo structures (Fig. S10) and homology models (65) of the m^{6}A-bound state confirmed a high degree of conservation of the key residues required for m^{6}A-RNA binding and overall folding of ECT1/9/11, in agreement with the partial functionality observed in our chimeric constructs (Fig. 3*B*). However, we also detected a few non-conservative amino acid substitutions specific for ECT1/9/11 (Figs. 4*A* and S9), preferentially located on the surface of the protein but far from the RNA-binding groove (Fig. S10). These non-conservative substitutions are prime candidates to explain the minor differences in functionality, a scenario that would imply biological functions of YTH domains other than mere m^{6}A-binding capacity.

The IDRs of ECT1/9/11 have biophysical properties distinct from ECT2/3/4/5/6/7/8/10

We next focused on the N-terminal regions of *ECT1/9/11.* Because protein disorder predictions using MobiDB (66) confirmed that they are all disordered (Fig. S11), we used approaches suitable for the study of IDRs. Prediction of short linear motifs (SLiMs) that may mediate interaction with effector proteins gave many possibilities that did not follow any clear pattern when compared between the different ECT classes (Fig. S11). Hence, we decided to study other properties known to be important in IDRs such as the propensity to form liquid condensates, the amino acid composition, and the distribution of charge. ECT2 can form condensates *in vitro* (16), and localizes *in vivo* to distinct foci upon stress (16, 26), identified as stress granules (26). Similar properties have been reported for human YTHDFs, which engage in liquid-liquid phase transition when concentrated on polyvalent m^{6}A-modified target RNA *in vitro* (67-70). To determine whether the distinct functional classes of ECTs differed in their propensity to phase separate, we employed a recently developed coarse-grained

model of IDRs with residue-specific parameters estimated from experimental observations of many IDRs (71) to simulate their tendency to remain in the condensed phase or disperse in solution. As control, we used the well-studied IDR of the human RNA-binding protein Fused in Sarcoma (FUS) (72-75). The results showed that the IDR of ECT2 has a clear propensity to remain in the condensed phase (Fig. 4*B-C* and Movie S1) with a highly favorable Gibb's free energy change for the transition from dilute aqueous solution to condensed phase (Fig. 4*D*), comparable to, albeit not as pronounced as for the IDR of FUS (Fig. 4*B-D*). While the IDRs of most ECT proteins behaved comparably to ECT2, the IDRs of ECT1 and ECT11 were clear outliers (Figs. 4*B-D* and S12*A*). Compared to all other ECTs, the IDR of ECT1 showed a much stronger propensity to phase separate, whereas for ECT11, the IDR dispersed into solution much more readily (Fig. 4*B-D* and Movies S2,3). Thus, the IDRs of ECT1 and ECT11 have markedly different biophysical properties from those of ECT2/3/4/5/6/7/8/10, perhaps contributing to their different functionality *in vivo*. In contrast, the IDR of ECT9 did not stand out from the rest in the simulations of phase separation propensity. We note that the different phase separation propensity among the ECTs is driven, in the main, by their different content of sticky residues as defined by the CALVADOS model (71) (Fig. S12*B*). We also note in this context that ECT1, ECT9 and ECT11 all had a higher content of Phe residues than the nearly Phe-depleted IDRs of ECT2/3/4/5/6/7/8/10 (Fig. S12*C*). For ECT1, whose Tyr and Trp enrichment is comparable to that of ECT2/3/4/5/6/7/8/10 (Fig. S12*C*), the additional enrichment of the similarly sticky Phe may contribute to its more pronounced phase separation propensity, but it is unclear what, if any, the functional consequence of Phe enrichment in the Tyr/Trp-poor IDRs of ECT9 and ECT11 may be.

We finally analyzed the charge distribution, as this may dictate properties such as compactness and shape of IDR conformations (76, 77). We found that the IDRs of ECT2/3/4/5/6/7/8/10 show a recurrent pattern with accumulation of net negative charge towards the N-terminus, and recovery to near-neutrality generally closer to the YTH domain (Figs. 4*E* and S12*D*). The IDRs of ECT1, ECT9 and ECT11 differ from this pattern, but in distinct ways: ECT1 does not show a pronounced patterning of the charged residues, while ECT9 increasingly accumulates negative charge (Figs. 4*E*) and ECT11 increasingly accumulates positive charge (Fig. S12*D*). It is an interesting possibility, therefore, that charge separation in the IDR is a requirement to fulfill the molecular functions of ECTs to accelerate growth in primordial cells.

The ability to complement ect2/3/4 does not simply follow affiliation with phylogenetic clades

The results of our complementation assays clearly show that the ability to rescue defective timing of leaf emergence does not follow the defined phylogenetic clades, because both high

and low/null complementation capacities are present in the same clades. For example, ECT2, ECT5 and ECT8 in clades DF-A, -B and -C respectively present high ECT2/3/4-like functionality, while ECT1, ECT9 and ECT11 in these same clades do not (Figs. 1*B* and *2C-H*). It is noteworthy, however, that there is a clear trend for non-complementors to be the most highly divergent proteins in each group, as revealed by the length of the branches in the phylogenetic tree. Indeed, the longest branches in both our tree (highlighted in red in Fig. 1*B*) and the one in the study by Scutenaire et al. (26) correspond to ECT1, ECT9, and ECT11, which also have the highest number of amino acid substitutions compared to the only YTHDF protein of the liverwort *Marchantia polymorpha* (*Mpo* YTHDF) (78) (Fig. 5*A*). These results suggest that the common ancestor of YTHDF proteins in land plants may have had a molecular function related to that of the modern arabidopsis ECT2/3/4, and that this function has been conserved in the least-divergent members of the DF-A/B/C/D clades. Subsequent duplication and neofunctionalization may have driven the rapid evolution of highly-divergent YTHDF proteins in some clades (ECT11, ECT1, and ECT9 in Arabidopsis) that have lost their primary function and perhaps fulfill different molecular roles in the plant m^6A pathway.

The molecular function of ECT2/3/4 was present in the first land plants

To test whether the common ancestor of plant YTHDF proteins indeed had a molecular function similar to that of the modern *Arabidopsis* ECT2/3/4, we subjected *Mpo* YTHDF (Fig. 2*A*) to our functional assay, as bryophytes are the earliest-diverging taxon of land plants (Fig. 1*A*, (47-52)). The result showed a partial, but clear capacity of two different splice forms of *Mpo* YTHDF to complement the delay in leaf emergence and morphology defects of *Arabidopsis te234* plants (Figs. 5*B, S3D* and S13*A,B*). Although the complementation scores were low (Fig. 5*B*), a thorough inspection of the fluorescence intensity in all primary transformants revealed that the transgene was only expressed very weakly in a reduced subpopulation of plants, possibly be due to the tendency to trigger silencing upon heterologous expression of a long cDNA from a distant plant species (79). We therefore introduced the best complementing longer isoform into *rdr6-12/te234* mutants. This approach yielded a much higher complementation score of 52% (Fig. 5*C*), and plants very similar to those expressing ECT2 in the same background (Figs. 5*D* and S6*B*). Of note, no transformants showed exacerbation of the background phenotype in any case. Gratifyingly, the simulations of the behavior of the IDR of *Mpo* YTHDF showed phase separation propensity and charge pattern similar to the ECT2/3/4/5/6/7/8/10 group, further substantiating that these IDR characteristics correlate with the ability to fulfill ECT2/3/4 function in leaf formation (Fig. 5*E-H*), Hence, our results demonstrate that the core molecular activity of *Ath* ECT2/3/4 was already present in the first land plants, and suggest

that stimulating proliferation of primed stem cells is the ancestral YTHDF molecular function in plants.

Heterologous expression of animal YTHDFs enhances the phenotype caused by loss of ECT2/3/4 in Arabidopsis

Finally, we subjected *Hs* YTHDF2 and *Drosophila melanogaster* (*Dm*) YTHDF (Fig. 2*A*) to our functional assay to test whether they could function molecularly like plant YTHDFs. While we did not observe complementation of the delayed leaf emergence of *te234* plants in any of the transformants (Fig. 5*B*), several *Hs* YTHDF2 lines showed more acute developmental defects compared to *te234* (Fig. S13*B*). Although these defects were not observed among *Dm* YTHDF lines, no expression could be detected in these plants either. Hence, we repeated the transformation in the *rdr6-12/te234* background, this time obtaining numerous transformants of both *Dm* YTHDF and *Hs* YTHDF2 with strong fluorescence (Figs. 5*C* and S6*B*). These plants showed aberrant morphology comparable to the effect caused by expression of ECT9 and C-9/2, with resemblance to arabidopsis mutants with severe depletion of m^6A (Figs. 5*D* and S6*B*). Hence, we conclude that expression of metazoan YTHDFs in arabidopsis *ECT2/3/4*-deficient plants exacerbates their developmental phenotype. This result is in agreement with the generally accepted idea that m^6A/YTHDF2 destabilize mRNA in mammals (9, 80-83) while m^6A/ECT2/3/4 do the opposite in plants (23, 44, 84-86). However, an analogous molecular mechanism between plant and animal and plant YTHDFs cannot be completely ruled out, because a different cellular context may be hampering the molecular activity of metazoan YTHDFs in plant cells, and simple competition for targets with remaining ECTs could be the cause of the enhanced phenotype. Nonetheless, we note that the IDRs of *Hs* YTHDF2 and *Dm* YTHDF had much weaker phase separation propensity compared to ECT2 (Fig. 5*E-G*), perhaps contributing to the lack of ECT2/3-like function *in planta*.

Validation of the functional assay by loss-of-function genetic analysis

The results of our functional analysis predict that knockout of genes encoding arabidopsis YTHDF proteins with ECT2/3/4-like functionality (ECT5/6/7/8/10) may exacerbate the delay in leaf emergence of *te234* mutants if endogenously expressed in primordial cells. The reverse should also be true: mutation of *ECT1, ECT9* or *ECT11* should not exacerbate the developmental defects of *te234* even if these proteins are naturally expressed in leaf primordia. We chose *ECT1* for validation of this prediction for three reasons. First, mRNA-seq data shows that expression levels and tissue specificity of *ECT1* are comparable to that of *ECT4*, whose activity is easily revealed as an enhancer mutant of *ect2/ect3* (16), while *ECT9/11* are generally very lowly expressed (Fig. S14, (87)). Second, *ECT1* expression is

strong in shoot and root apices as judged by histochemical β-glucuronidase (GUS) assays of arabidopsis plants expressing *ECT1p:GUS* (36). Third, ECT1 is of particular interest because, as the closest relative of ECT2/3/4 in arabidopsis, the natural expectation would be some degree of redundancy. Indeed, the considerable expression of *ECT1* in the embryo (Fig. S14, (87)) may suggest that simultaneous mutation of *ECT1/ECT2/ECT3/ECT4* could phenocopy the embryonic lethality of mutants devoid of a functional methyltransferase complex (1, 44, 45). In this context, the lack of functional redundancy and therefore negligible genetic interaction predicted by our functional assay is not trivial, and hence appropriate for validation.

ECT1 and ECT2/3/4 are expressed in the same cell types

We first assessed the expression pattern of ECT1 protein using stable transgenic lines expressing a C-terminal TFP-fusion of *ECT1* (*ECT1p:gECT1-FLAG-TFP-ECT1ter*, henceforth *ECT1-TFP*, Fig. S15*A*). The pattern of turquoise fluorescence was strongly reminiscent of that of fluorescent ECT2/3/4 fusions (16) with signal at the base of young leaves, in leaf primordia, main root tips and lateral root primordia at different stages (Fig. 6*A-D*). Thus, the overlap in ECT1 and ECT2/3/4 expression patterns is substantial. We note that these transgenic lines selected on the basis of visible and reproducible pattern of fluorescence may have elevated expression levels compared to the endogenous *ECT1*, because we could detect *ECT1-TFP* mRNA but not the endogenous transcript by RNA blotting (Fig. S15*B*).

ECT1 resides in the cytoplasm where it can form granules

To perform ECT2/3/4-like functions, ECT1 would also need to be cytoplasmic. Confocal microscopy of meristematic root cells indeed showed cytoplasmic signal with heterogenous texture similar to what we previously described for ECT2/3/4 (16) (Fig. 6*E*). However, most root tips had a few cells containing distinct ECT1-TFP foci (Fig. 6*F*) that formed in the absence of stress. Importantly, we found no trace of free TFP that could be causing artifacts in the imaging (Fig. S15*A*). We conclude from the similar expression pattern and subcellular localization of ECT1 and ECT2/3/4 that assessment of the possible exacerbation of *te234* phenotypes by additional loss of *ECT1* function is a meaningful test of our functional assay. We also note that the spontaneous granule formation of ECT1-TFP *in vivo* is consistent with the strong phase separation propensity revealed by the molecular simulations (Fig. 5*B-D* and Movie S2).

Mutation of ECT1 does not exacerbate the phenotype of ect2/ect3/ect4 plants

We isolated three lines homozygous for T-DNA insertions in the *ECT1* gene body, *ect1-1* (SAIL_319_A08), *ect1-2* (GK-547H06), and *ect1-3* (SALK_059722) (Fig. 6*G*) and, based on qPCR analyses, concluded that the three are knockout mutations (Fig. 6*H*). We used the *ect1-1* and *ect1-2* alleles for subsequent systematic genetic analyses with a series of single, double, triple and quadruple mutants of *ect1* combined with *ect2*, *ect3* and *ect4*. Importantly, both *ect1-1/ect2-1/ect3-1/ect4-2 (qe1234*, for 'quadruple *ect1/2/3/4*', see Fig. S15*C* for multiple mutant abbreviations) and *ect1-2/ect2-3/ect3-2/ect4-2* (*Gqe1234*, for 'GABI-KAT quadruple *ect1/2/3/4*') mutant plants were viable. Furthermore, the timing of the leaf emergence and the overall rosette morphology were indistinguishable from the corresponding *te234* and *Gte234* parental mutants (Fig. 6*I* and S15*D*). Similarly, *ect1-2/ect2-3/ect3-2* (*Gte123*) mutants were identical to *ect2-3/ect3-2 (Gde23)*, and single *ect1* mutants did not show any obvious defects, including trichome branching and root morphology (Figs. 6*I* and S15*D-F*). Thus, there is no indication of redundancy between ECT1 and the other three DF-A clade proteins, despite the fact that ECT1 is expressed in a pattern similar to that of the other DF-A paralogs. This conclusion agrees with the poor ECT2/3/4-like activity of ECT1 in our functional assays, thereby validating the approach.

Discussion

Functional versus biological redundancy

Our study reveals that most YTHDF proteins in higher plants can perform, essentially, the same molecular functions that allow ECT2/3/4 to promote stem cell proliferation and control organogenesis. We anticipate that this molecular function may be used in specific biological contexts by ECT5/6/7/8/10 due to different expression patterns and response to stimuli, thereby achieving biological specialization with respect to ECT2/3/4. Such variability can already be inferred from available RNA-Seq data in the arabidopsis eFP Browser (BAR.toronto.ca) that show, for example, that ECT8 is primarily expressed during senescence-associated processes, perhaps in connection to re-growth after reversible proliferative arrest (88, 89). In addition to redundancy, there is a clear pattern of lineage-specific specialization in fast-evolving YTHDF proteins from different clades, exemplified in arabidopsis by ECT1, ECT11 and ECT9. Whether these proteins act to counteract the ECT2/3/4-like function by competing for m^6A-containing transcripts, perhaps halting growth upon stress, or whether they perform independent functions in different tissues or cell types, is open for new investigations.

ECT1 is a newly evolved gene in a Brassicaceae lineage

Surprisingly, the only member of the DF-A clade not yet characterized as m^6A reader, ECT1, scored low in our functional assay and indeed, inactivation of *ECT1* does not enhance the

phenotype caused by loss of *ECT2/3/4* despite expression of *ECT1* in the same cell types and subcellular compartment as the other DF-A paralogs. We note that a previous description of nuclear and cytoplasmic fluorescence of a *35Sp:ECT1-GFP* fusion (36) may be due to the fact that the authors used epifluorescence instead of confocal fluorescence microscopy for the analyses. Because that technique does not distinguish between inside or around the nucleus, the signal observed may have been perinuclear, e.g. from the endoplasmic reticulum. A thorough look at phylogenetic analyses from Scutenaire *et al.* (26) reveals that *ECT1* is only present in species of the 'lineage I' of Brassicaceae such as *A. thaliana* and *Capsella rubella* (*Cru*) (90). The closely related *Brassica oleracea* ('lineage II') (90) has a clear homolog of *Ath/Cru ECT3*, but not *ECT1* (26), while other dicots outside *Brassicaceae* encode *ECT3/1* orthologs in a separate cluster (26). Therefore, *ECT1* is the product of a recent duplication and neofunctionalization event of the *ECT3/1* gene in a lineage of *Brassicaceae* that diverged more rapidly than any other DF-A clade member in *Arabidopsis* (Fig. 1*B*, (26)). This agrees with our overall conclusion that most YTHDFs in early angiosperms possessed *Mpo* YTHDF/*Ath* ECT2/3/4 molecular functions, and only late events resulted in specialization of a few YTHDF proteins. For ECT1, this occurred late in a small subgroup of dicots. We note that such rapid neofunctionalization appears to result primarily from changes in the IDR. We were able to identify two biophysical properties of the IDR of ECT1, a stronger predicted propensity to phase-separate and its charge distribution, that are distinct from the ECT2/3/4/5/6/7/8/10 group, but note that additional changes, for example in SLiMs mediating interaction with regulatory factors, may also contribute to the different functionality.

The diversity of plant YTHDF proteins and the need for many YTHDFs

The observation that a few fern YTHDF proteins cluster in the angiosperm clade DF-D points to an early diversification of plant YTHDFs, while the fact that some fern and gymnosperm YTHDFs define an independent clade (DF-F) suggests that complex plants may need many different YTHDFs, well surpassing the three paralogs found in most vertebrates (6) or in any other types of eukaryotes (8, 57). But why? The need for many YTHDFs of considerable variability is all the more puzzling given our demonstration here that the majority of them retains the same molecular functions. Robustness is a commonly given explanation for redundancy, but the same argument could apply to methyltransferase subunits, and *Arabidopsis* has maintained only one copy of *MTA*, *MTB*, *FIP37*, *VIR* and *HAKAI* (57). Therefore, the answer should probably be sought elsewhere. It is clear now that the m^6A/YTHDF axis controls cellular proliferation and organ growth in plants (17, 91), and a principal difference between animals and plants is the extreme plasticity of plant growth. Because growth and stress responses are also tightly intertwined, it is possible that plants

use a plethora m^6A-YTHDF growth-stimulators in different organs, responsive to different stresses and developmental cues, to shape plant architecture in response to stimuli. The answer to the paradox may be in the details: plant YTHDFs responsive to different environmental cues may fine-tune growth programs through slightly different performance and target-binding capacity in partly overlapping expression domains, where they are known to compete for targets (23).

Materials and Methods

Plant material and growth conditions

All the used lines are in the *A. thaliana* Columbia-0 (Col-0) ecotype. The T-DNA insertion lines *rdr6-12* (61), *ect1-1* (SAIL_319_A08), *ect1-2* (GK-547H06), and *ect1-3* (SALK_059722) were obtained from Nottingham Arabidopsis Stock Centre. The mutants *ect2-3*, *ect3-2*, *ect2-3/ect3-2* (*Gde23*), *ect2-3/ect3-2/ect4-2* (*Gte234*) and *ect2-1/ect3-1/ect4-2* (*te234*) used for genetic crosses or as background to produce transgenic lines, and the primers for genotyping the corresponding alleles, have been previously described (16, 17). The standard growing conditions used are detailed in *SI Materials and Methods*.

Phylogenetic analysis

Amino acid sequences of selected YTH-domain proteins were downloaded from UniProt (Apweiler et al., 2004), TAIR (www.arabidopsis.org), phytozome (92) or GinkgoDB (93) and trimmed from their N-termini to discard intrinsically disordered regions with poor sequence conservation, thereby retaining only the regions with conservation to those structured in human YTHDF1 (62) (*SI Protein Sequences*, Dataset S1). Trimmed sequences were aligned with MUSCLE (94, 95) with default parameters (Dataset S2). Evolutionary analyses were conducted using IQ-Tree 2 (96, 97) from the CIPRES Science Gateway v. 3.3 (https://www.phylo.org/index.php/) to infer a phylogenetic tree using maximum likelihood. The model based on revised JTT Matrix (98) JTTDCMut+R5 was determined as best-fit using the built-in tool ModelFinder (99) according to Bayesian Information Criterion (BIC) scores, and used for the analysis. *Mpo* YTHDF was selected as outgroup and drawn as root. Measures of branch supports by approximate Likelihood Ratio Test [aLRT] (100) (SH-aLRT, %) were calculated, and the resulting values were used for color-coding the nodes with the tool FigTree (https://github.com/rambaut/figtree/). The precise SH-aLRT values of all nodes in newick format and additional information can be found in Dataset S3.

Estimates of evolutionary divergence between sequences were obtained computing pairwise distances among the full length sequences of the same proteins (Dataset S4), aligned in the same manner (Dataset S5) using the Equal Input model (101) in MEGA11 (102). The

positions containing gaps and missing data were eliminated (complete deletion option), resulting a total of 180 positions in the final dataset.

Cloning, plant transformation and line selection

We employed the scar-free USER cloning method (103) to produce *uS7Bp:cYTHDF(X)-mCherry-OCSter* and *ECT1p:gECT1-FLAG-TFP-ECT1ter* constructs by gluing PCR-amplified DNA fragments into the pCAMBIA3300-U plasmid (104) that provides resistance to glufosinate in plants. Cloning of the trial *ECT2p:gECT1-mCherry-ECT2ter* and *ECT2p:gECT4-mCherry-ECT2ter* constructs was done by GreenGate (105). The detailed cloning, transformation and line selection strategy for each construct, as well as the specific details of T1 selections for the complementation assay, are explained in *SI Materials and Methods*.

IDR modeling and simulations

We carried out coarse-grained simulations of protein IDRs using the implicit-solvent CALVADOS 2 forcefield (71, 106). In CALVADOS, amino-acid residues are mapped to single beads corresponding to their amino-acid type. Neighboring residues are linked using harmonic bonds with an equilibrium distance of 0.38 nm and force constant of 8033 kJ $mol^{-1}nm^{-2}$. Nonbonded interactions are modeled with a hydropathy amino-acid "stickiness" model using λ-parameters from CALVADOS 2 (106). Salt-screened electrostatic interactions are accounted for with a truncated (r_c = 4 nm) and shifted Debye-Hückel potential; non-electrostatic interactions were truncated at 2 nm (106). We used a Langevin integrator with friction coefficient γ = 0.01 ps^{-1} to propagate the simulation in 0.01 ps timesteps. All simulations were carried out at T = 293 K.

We performed slab simulations of IDRs of the following proteins: ECT1-ECT11, *Mpo* YTHDF, *Hs* YTHDF2, *Dm* YTHDF1, and *Hs* FUS, using a previously described procedure (106, 107). IDR sequences are detailed in *SI Protein Sequences*. Briefly, we simulated 100 copies of each IDR (one type of IDR at a time) in a simulation slab of dimensions 25 nm x 25 nm x 300 nm (ECT2, ECT4, ECT5, *Mpo* YTHDF, *Hs* YTHDF2, *Dm* YTHDF), 17 nm x 17 nm x 300 nm (ECT1, ECT3, ECT6-ECT11), or 15 nm x 15 nm x 150 nm (*Hs* FUS), accounting for the different IDR sequence lengths. We simulated each system for 7500 ns and recorded configurations at 1 ns intervals. The first 2000 ns of each simulation were considered as equilibration and not used for analysis. We centered the protein condensates in the z-dimension of the simulation box during post-processing. We determined concentrations of the dense phase and dilute phase by fitting the concentration profiles to a hyperbolic function, as described by Tesei et al. (71). Excess transfer free energies were calculated

from the dilute and dense phase equilibrium concentrations via $\Delta G_{trans} = RT \ln (c_{dilute} / c_{dense})$ (108).

Charges (q) were added along the IDR sequences to obtain cumulative charge plots. Arg and Lys residues were assigned charges of q = +1; Glu and Asp residues were assigned charges of q = -1; other residue types, including His, were assigned charges of q = 0. The average 'stickiness' of a given IDR sequence was determined by calculating the mean hydrophobicity $\bar{\lambda}$ of the IDR amino acid composition, using the CALVADOS 2 hydrophobicity scale reported in (106).

Additional standard methods including DNA and cDNA obtention, genotyping, fluorescence microscopy, qPCR, RNA and protein blotting, and the characterization of YTH domains, are described in *SI Materials and Methods*.

Acknowledgements and funding sources

We thank Lena Bjørn Johansson, Sergio D'Anna and Tobias Lahti for technical assistance, Theo Bølsterli and his team for plant care, and the Danish National Life Science Supercomputing Center Computerome and the core facility for biocomputing at the Department of Biology (University of Copenhagen) for the use of computational resources. We acknowledge the constructive criticism of Rupert Fray as a reviewer on previous publications, as his suggestion to use the *ect2/ect3/ect4* triple mutant phenotype to test ECT molecular function by systematic ectopic expression in the ECT2 expression domain inspired us to embark on the present project. This work was supported by the Novo Nordisk Foundation through a Hallas-Møller Ascending Investigator 2019 grant NNF19OC0054973 to P.B. and an NNF18OC0033950 grant to K.L.-L. via the PRISM (Protein Interactions and Stability in Medicine and Genomics) center, and by the Independent Research Fund Denmark through a Research Project2 grant (9040-00409B) from to P.B.

References

1. S. Zhong *et al.*, MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* **20**, 1278-1288 (2008).

2. S. Geula *et al.*, m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* **347**, 1002-1006 (2015).

3. N. Liu *et al.*, N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560-564 (2015).

4. C. Roost *et al.*, Structure and Thermodynamics of N6-Methyladenosine in RNA: A Spring-Loaded Base Modification. *J. Am. Chem. Soc.* **137**, 2107-2115 (2015).

5. R. C. Spitale *et al.*, Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **519**, 486-490 (2015).

6. D. P. Patil, B. F. Pickering, S. R. Jaffrey, Reading m6A in the Transcriptome: m6A-Binding Proteins. *Trends Cell Biol.* **28**, 113-127 (2018).

7. D. Dominissini *et al.*, Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206 (2012).

8. P. Stoilov, I. Rafalska, S. Stamm, YTH: a new domain in nuclear proteins. *Trends Biochem. Sci.* **27**, 495-497 (2002).

9. X. Wang *et al.*, N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117-120 (2014).

10. F. Li, D. Zhao, J. Wu, Y. Shi, Structure of the YTH domain of human YTHDF2 in complex with an m6A mononucleotide reveals an aromatic cage for m6A recognition. *Cell Res.* **24**, 1490-1492 (2014).

11. S. Luo, L. Tong, Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc. Natl. Acad. Sci. USA* **111**, 13834-13839 (2014).

12. D. Theler, C. Dominguez, M. Blatter, J. Boudet, F. H. Allain, Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA. *Nucleic Acids Res.* **42**, 13911-13919 (2014).

13. C. Xu *et al.*, Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* **10**, 927-929 (2014).

14. T. Zhu *et al.*, Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. *Cell Res.* **24**, 1493-1496 (2014).

15. Z. Zhang *et al.*, The YTH domain is a novel RNA binding domain. *The Journal of biological chemistry* **285**, 14701-14710 (2010).

16. L. Arribas-Hernández *et al.*, An m6A-YTH Module Controls Developmental Timing and Morphogenesis in Arabidopsis. *Plant Cell* **30**, 952-967 (2018).

17. L. Arribas-Hernández *et al.*, Recurrent requirement for the m6A-ECT2/ECT3/ECT4 axis in the control of cell proliferation during plant organogenesis. *Development* **147**, dev189134 (2020).

18. L. Lasman *et al.*, Context-dependent functional compensation between Ythdf m6A reader proteins. *Genes Dev.* **34**, 1373-1391 (2020).

19. I. Ivanova *et al.*, The RNA m6A reader YTHDF2 is essential for the post-transcriptional regulation of the maternal transcriptome and oocyte competence. *Mol. Cell* **67**, 1059-1067 (2017).

20. M. Li *et al.*, Ythdf2-mediated m6A mRNA clearance modulates neural development in mice. *Genome biology* **19**, 69 (2018).

21. H. Shi *et al.*, m6A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* **563**, 249-253 (2018).

22. C. Kontur, M. Jeong, D. Cifuentes, A. J. Giraldez, Ythdf m6A Readers Function Redundantly during Zebrafish Development. *Cell Rep.* **33**, 108598 (2020).

23. L. Arribas-Hernández *et al.*, The YTHDF proteins ECT2 and ECT3 bind largely overlapping target sets and influence target mRNA abundance, not alternative polyadenylation. *eLife* **10**, e72377 (2021).

24. T. Lence *et al.*, m6A modulates neuronal functions and sex determination in Drosophila. *Nature* **540**, 242 (2016).

25. I. U. Haussmann *et al.*, m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. *Nature* **540**, 301 (2016).

26. J. Scutenaire *et al.*, The YTH Domain Protein ECT2 Is an m6A Reader Required for Normal Trichome Branching in Arabidopsis. *Plant Cell* **30**, 986-1005 (2018).

27. D. Li *et al.*, Genome-wide identification, biochemical characterization, and expression analyses of the YTH domain-containing RNA-binding protein family in Arabidopsis and rice. *Plant Mol. Biol. Report.* **32**, 1169-1186 (2014).

28. H. Du *et al.*, YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* **7**, 12626 (2016).

29. O. H. Park *et al.*, Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol. Cell* **74**, 494-507.e498 (2019).

30. S. H. Boo *et al.*, UPF1 promotes rapid degradation of m6A-containing RNAs. *Cell Rep.* **39**, 110861 (2022).

31. L. Arribas-Hernández *et al.*, Principles of mRNA targeting via the Arabidopsis m6A-binding protein ECT2. *eLife* **10**, e72375 (2021).

32. S. Brodsky *et al.*, Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. *Mol. Cell* **79**, 459-471.e454 (2020).

33. J. Scutenaire *et al.*, The S. cerevisiae m6A-reader Pho92 promotes timely meiotic recombination by controlling key methylated transcripts. *Nucleic Acids Res.* 10.1093/nar/gkac640, gkac640 (2022).

34. R. A. Varier *et al.*, N6-methyladenosine (m6A) reader Pho92 is recruited co-transcriptionally and couples translation to mRNA decay to promote meiotic fitness in yeast. *eLife* **11**, e84034 (2022).

35. S. Zaccara, S. R. Jaffrey, A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. *Cell* **181**, 1582-1595.e1518 (2020).

36. S. H. Ok *et al.*, Novel CIPK1-associated proteins in Arabidopsis contain an evolutionarily conserved C-terminal region that mediates nuclear localization. *Plant Physiol.* **139**, 138-150 (2005).

37. L.-P. Bergeron-Sandoval, N. Safaee, Stephen W. Michnick, Mechanisms and Consequences of Macromolecular Phase Separation. *Cell* **165**, 1067-1079 (2016).

38. W. Borcherds, A. Bremer, M. B. Borgia, T. Mittag, How do intrinsically disordered protein regions encode a driving force for liquid–liquid phase separation? *Curr. Opin. Struct. Biol.* **67**, 41-50 (2021).

39. X. Wang *et al.*, N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388-1399 (2015).

40. H. Shi *et al.*, YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* **27**, 315-328 (2017).

41. A. Li *et al.*, Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. *Cell Res.* **27**, 444 (2017).

42. R. G. Fray, G. G. Simpson, The Arabidopsis epitranscriptome. *Curr. Opin. Plant Biol.* **27**, 17-21 (2015).

43. Z. Bodi *et al.*, Adenosine methylation in Arabidopsis mRNA is associated with the 3' end and reduced levels cause developmental defects. *Front. Plant Sci.* **3**, 48 (2012).

44. L. Shen *et al.*, N6-methyladenosine RNA modification regulates shoot stem cell fate in Arabidopsis. *Dev. Cell* **38**, 186-200 (2016).

45. K. Růžička *et al.*, Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol.* **215**, 157-172 (2017).

46. N. D. Pires, L. Dolan, Morphological evolution in land plants: new designs with old genes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 508-518 (2012).

47. I. One Thousand Plant Transcriptomes, One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679-685 (2019).

48. B. J. Harris, C. J. Harrison, A. M. Hetherington, T. A. Williams, Phylogenomic Evidence for the Monophyly of Bryophytes and the Reductive Evolution of Stomata. *Current Biology* **30**, 2001-2012.e2002 (2020).

49. F.-W. Li *et al.*, Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nature plants* **6**, 259-272 (2020).

50. M. N. Puttick *et al.*, The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte. *Current Biology* **28**, 733-745.e732 (2018).

51. D. Su *et al.*, Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants. *Mol. Biol. Evol.* **38**, 3332-3344 (2021).

52. J. Zhang *et al.*, The hornwort genome and early land plant evolution. *Nature plants* **6**, 107-118 (2020).

53. S. Magallón, K. W. Hilu, D. Quandt, Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am. J. Bot.* **100**, 556-573 (2013).

54. T. Lan, W. Xiong, X. Chen, B. Mo, G. Tang, Plant cytoplasmic ribosomal proteins: an update on classification, nomenclature, evolution and resources. *Plant J.* **110**, 292-318 (2022).

55. D. Weijers *et al.*, An Arabidopsis Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* **128**, 4289-4299 (2001).

56. M. Fagard, H. Vaucheret, (TRANS)GENE SILENCING IN PLANTS: How Many Mechanisms? *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**, 167-194 (2000).

57. D. L. Balacco, M. Soller, The m6A Writer: Rise of a Machine for Growing Tasks. *Biochemistry* **58**, 363-378 (2019).

58. C. Wang *et al.*, A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. *Nucleic Acids Res.* **44**, 969-982 (2016).

59. T. Dalmay, A. Hamilton, S. Rudd, S. Angell, D. C. Baulcombe, An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**, 543-553 (2000).

60. P. Mourrain *et al.*, *Arabidopsis SGS2* and *SGS3* genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**, 533-542 (2000).

61. A. Peragine, M. Yoshikawa, G. Wu, H. L. Albrecht, R. S. Poethig, SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes Dev.* **18**, 2368-2379 (2004).

62. C. Xu *et al.*, Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *Journal of Biological Chemistry* **290**, 24902-24913 (2015).

63. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

64. M. Varadi *et al.*, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439-D444 (2022).

65. M. Biasini *et al.*, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252-W258 (2014).

66. D. Piovesan *et al.*, MobiDB: 10 years of intrinsically disordered proteins. *Nucleic Acids Res.* **51**, D438-D444 (2023).

67. J. Wang *et al.*, Binding to m6A RNA promotes YTHDF2-mediated phase separation. *Protein & Cell* **11**, 304-307 (2020).

68. Y. Gao *et al.*, Multivalent m6A motifs promote phase separation of YTHDF proteins. *Cell Res.* **29**, 767-769 (2019).

69. R. J. Ries *et al.*, m6A enhances the phase separation potential of mRNA. *Nature* **571**, 424-428 (2019).

70. Y. Fu, X. Zhuang, m6A-binding YTHDF proteins promote stress granule formation. *Nat. Chem. Biol.* **16**, 955-963 (2020).

71. G. Tesei, T. K. Schulze, R. Crehuet, K. Lindorff-Larsen, Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proc. Natl. Acad. Sci. USA* **118**, e2111696118 (2021).

72. M. Kato, S. L. McKnight, The low-complexity domain of the FUS RNA binding protein self-assembles via the mutually exclusive use of two distinct cross-β cores. *Proc. Natl. Acad. Sci. USA* **118**, e2114412118 (2021).

73. M. Kato *et al.*, Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753-767 (2012).

74. A. Patel *et al.*, A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066-1077 (2015).

75. Jacob C. Schwartz, X. Wang, Elaine R. Podell, Thomas R. Cech, RNA Seeds Higher-Order Assembly of FUS Protein. *Cell Rep.* **5**, 918-925 (2013).

76. G. Bianchi *et al.*, Distribution of Charged Residues Affects the Average Size and Shape of Intrinsically Disordered Proteins. http://dx.doi.org/10.3390/biom12040561.

77. R. K. Das, Y. Huang, A. H. Phillips, R. W. Kriwacki, R. V. Pappu, Cryptic sequence features within the disordered protein p27Kip1 regulate cell cycle signaling. *Proc. Natl. Acad. Sci. USA* **113**, 5616-5621 (2016).

78. J. L. Bowman *et al.*, Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell* **171**, 287-304.e215 (2017).

79. L. Wang, M. J. Roossinck, Comparative analysis of expressed sequences reveals a conserved pattern of optimal codon usage in plants. *Plant Mol. Biol.* **61**, 699-710 (2006).

80. S. Sommer, U. Lavi, J. E. Darnell, The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *Journal of Molecular Biology* **124**, 487-499 (1978).

81. Y. Wang *et al.*, N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nature cell biology* **16**, 191-198 (2014).

82. S. Ke *et al.*, m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.* **31**, 990-1006 (2017).

83. V. A. Herzog *et al.*, Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Meth.* **14**, 1198 (2017).

84. S. J. Anderson *et al.*, N(6)-Methyladenosine Inhibits Local Ribonucleolytic Cleavage to Stabilize mRNAs in Arabidopsis. *Cell Rep.* **25**, 1146-1157 (2018).

85. L.-H. Wei *et al.*, The m6A Reader ECT2 Controls Trichome Morphology by Affecting mRNA Stability in Arabidopsis. *Plant Cell* **30**, 968-985 (2018).

86. M. T. Parker *et al.*, Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m6A modification. *eLife* **9**, e49658 (2020).

87. H. Zhang *et al.*, A Comprehensive Online Database for Exploring ~20,000 Public Arabidopsis RNA-Seq Libraries. *Molecular Plant* **13**, 1231-1233 (2020).

88. S. Gan, Mitotic and Postmitotic Senescence in Plants. *Science of Aging Knowledge Environment* **2003**, re7-re7 (2003).

89. S. E. Wuest, M. A. Philipp, D. Guthörl, B. Schmid, U. Grossniklaus, Seed Production Affects Maternal Growth and Senescence in Arabidopsis *Plant Physiol.* **171**, 392-404 (2016).

90. L. A. Nikolov *et al.*, Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. *New Phytol.* **222**, 1638-1651 (2019).

91. L. Arribas-Hernández, P. Brodersen, Occurrence and functions of m6A and other covalent modifications in plant mRNA. *Plant Physiol.* **182**, 79-96 (2020).

92. D. M. Goodstein *et al.*, Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178-D1186 (2012).

93. K.-J. Gu, C.-F. Lin, J.-J. Wu, Y.-P. Zhao, GinkgoDB: an ecological genome database for the living fossil, Ginkgo biloba. *Database* **2022**, baac046 (2022).

94. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797 (2004).

95. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).

96. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **32**, 268-274 (2015).

97. B. Q. Minh *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530-1534 (2020).

98. C. Kosiol, N. Goldman, Different Versions of the Dayhoff Rate Matrix. *Mol. Biol. Evol.* **22**, 193-199 (2005).

99. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Meth.* **14**, 587-589 (2017).

100. H. Shimodaira, M. Hasegawa, Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. *Mol. Biol. Evol.* **16**, 1114-1114 (1999).

101. F. Tajima, M. Nei, Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* **1**, 269-285 (1984).

102. K. Tamura, G. Stecher, S. Kumar, MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **38**, 3022-3027 (2021).

103. J. Bitinaite, N. M. Nichols, DNA cloning and engineering by uracil excision. *Curr. Protoc. Mol. Biol.* **Chapter 3**, Unit 3 21 (2009).

104. H. H. Nour-Eldin, B. G. Hansen, M. H. Norholm, J. K. Jensen, B. A. Halkier, Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res.* **34**, e122 (2006).

105. A. Lampropoulos *et al.*, GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis. *PLOS ONE* **8**, e83043 (2013).

106. G. Tesei, K. Lindorff-Larsen (2022) Improved Predictions of Phase Behaviour of Intrinsically Disordered Proteins by Tuning the Interaction Range. (bioRxiv).

107. G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, J. Mittal, Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Computational Biology* **14**, e1005941 (2018).

108. Z. Benayad, S. von Bülow, L. S. Stelzl, G. Hummer, Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *Journal of Chemical Theory and Computation* **17**, 525-537 (2021).

Figure Legends

Fig. 1. Phylogenetic analysis of YTHDF proteins in land plants. (*A*) Schematic representation of land plant evolution and relative position of the species used in this study.

The architecture of the diagram and the age (million years, My) indicated on some nodes are a simplified version of the tree from Su et al. (51). The length of the branches has been adjusted for illustrative purposes. (*B*) Phylogenetic tree of YTHDF proteins in land plants based on the sequence alignment of regions with homology to the structured regions of the metazoan orthologs (62) (Datasets S1, S2). *Marchantia polymorpha* (Mpo) YTHDF is chosen as outgroup. *Arabidopsis thaliana* (*Ath*) YTHDF proteins are named after the original nomenclature for proteins containing an Evolutionarily Conserved C-Terminal Region (ECT) established by Ok et al. (36). Proteins from other plant species adhere to the nomenclature established by Scutenaire et al. (26), with small variations reflecting additional clades (DF-E, -F) and the split of the original DF-C clade into -C and -D. Statistical support is calculated as approximate likelihood-ratio test (aLRT) and indicated by colour-coded spheres on all nodes, but the most relevant values are specified. The length of the branches represents the evolutionary distance in number of amino acid substitutions per site relative to the scale bar. Color-coding and abbreviations of species as in A. Dotted lines highlight weakly supported affiliations to clades.

Fig. 2. Most but not all arabidopsis YTHDF proteins can replace ECT2/3/4-function. (*A*) Diagram showing the relative length of the N-terminal IDRs of *Arabidopsis thaliana* (*Ath*) YTHDF proteins (ECT1-ECT11) together with *Marchantia polymorpha* (*Mpo*) YTHDF, *Homo sapiens* (*Hs*) YTHDF2, and *Drosphila melanogaster* (*Dm*) YTHDF. Numbers indicate the length of the proteins in amino acids (aa). (*B*) Strategy followed for the functional assay. *uS7Bp:cECT(X)-mCherry* constructs are introduced in *ect2-1/ect3-1/ect4-2* (*te234*) plants, and complementation rates are estimated by the percentage of primary transformants (T1) whose first true leaves have a size (s) exceeding 0.5 mm after 10 days of growth. The *construct uS7Bp:mCherry* is used as negative control. Examples of T1 seedlings expressing control and ECT2 constructs are shown. Dashed outlines are magnified below to show mCherry fluorescence in emerging leaves. (*C*) Weighed averages of the complementation percentages observed for each *uS7Bp:cECT(X)-mCherry* construct in 2-5 independent transformations (I.T.). Letters between parenthesis indicate the DF- clade assigned to each protein. Coloring highlights ECT2 (reference) and proteins with low or negligible complementation capacity (ECT1/9/11) for fast identification along the manuscript. Low transformation efficiency (T.E.) for ECT9 can be infered by the low number of transformants (n) highlighted in red, but detailed T.E. and raw complementation rates in each transformation can be found in Figs. S5 and S3A respectively. Additional data regarding this assay is shown in Fig. S4. (*D*) Same as C on a single transformation using the *rdr6-12/te234* background. (*E-H*) Rosette phenotype of 32-day old T1s bearing *uS7Bp:cECT(X)-mCherry* constructs in the *rdr6-12/te234* background, sorted according to their degree of complementation in three panels: *F* (high), *G* (residual) and *H* (non-complementation and enhancement). Different genetic backgrounds and control transformants (mCherry) are shown in *E* as a reference. The scalebar (5 cm) applies to all plants in the four panels.

Fig. 3. Dissection of ECT1/9/11 functional regions. (*A*) Schematic representation of the strategy followed to express chimeras with N-terminal IDR and YTH domains of different ECT proteins. (*B*) Weighed averages of the complementation rates observed for each chimeric construct measured as in Fig. 2B,C. The raw complementation rates and transformation efficiencies of each independent transformation (I.T.) can be found in Figs. S3B and S5 respectively, and photographs of the scored 10-day-old T1 seedlings with mCherry fluorescence are shown in Fig. S8. (*C*) 10-day-old primary transformants of *rdr6-12/te234* with the indicated transgenes. Dashed outlines are magnified below to show mCherry fluorescencece. (*D*) Same genotypes as in *C* at 26 days after germination. Additional independent transgenic lines, developmental stages and controls for *C* and *D* can be found in Fig. S6. All transgenes are expressed from the *uS7B* promoter and contain mCherry fused to the C-terminus.

Fig. 4. Analysis of ECT1/9/11 functional regions. (*A*) Topology of the YTH domain of *Homo sapiens* (*Hs*) YTHDF1 with secondary structure elements drawn in pink (62). The two grey helices represent the only structural elements outside the YTH domain (as defined by Stoilov et al. (8)): α0 immediately upstream, and α4 at the C-terminus of the protein. Superimposed on the diagrams are the amino acids in ECT1, ECT9 and ECT11 that constitute non-conservative substitutions compared to the majority of ECTs (see alignment in Fig. S9). The localization of these amino acids in the predicted 3D structures, predominantly on the surface, and the spatial integrity of the aromatic cage in the homology models of ECT1/9/11 can be found in Fig. S10. (*B*) Slab simulations of the IDRs of the indicated proteins showing their tendency to remain in condensed phase or disperse in solution. (*C*) Time-averaged density profiles for a selected group of the slab simulations shown in *B*. The remaining proteins are shown in Fig. S12A. (*D*) Calculated free energy change for the transition between aqueous solution and condensed phase ($\Delta G_{trans} = RT \ln (c_{dilute} / c_{dense})$) for the IDRs of the indicated proteins. (*E*) Cumulative charge distribution of the IDRs of the indicated proteins. The length in amino acids (aa) is at scale. The same analysis for all ECTs can be found in Fig. S12D.

Fig. 5. Heterologous expression of *Marchantia polymorpha* (*Mpo*) YTHDF in arabidopsis rescues loss of ECT2/3/4 function, while *Homo sapiens* (*Hs*) or *Drosophila melanogaster* (*Dm*) YTHDF proteins aggravates it. (*A*) Evolutionary divergence among ECT proteins estimated by pairwise comparison of the proteins analyzed in Fig. 1B. (*B*) Weighed averages of the complementation rates observed for each of the indicated constructs (cDNA fused to mCherry under the control of the *uS7B* promoter) as in Fig. 2C in the te234 background. S (short) and L (long) refer to the two different YTHDF spliceforms found among *M. polymorpha* cDNAs. Transformation efficiencies and raw complementation rates in each transformation can be found in Figs. S5 and S3 respectively. (*C*) Same as in *B* but from one transformation in the *rdr6-12/te234* background. (*D*) 32-day-old primary transformants of *rdr6-12/te234* expressing the heterologous transgenes. Additional controls and developmental stages are shown in Fig. S6. (*E*) Slab simulations of the IDRs of the indicated proteins as in Fig. 4B. (*F*) Time-averaged density profiles for the slab simulations in *E*. (*G*) Calculated free energy change for the IDRs of the indicated proteins as in Fig. 4D. (*H*) Cumulative charge distribution of the IDRs of the indicated proteins as in Fig. 4E. The same analysis for all arabidopsis ECTs and animal YTHDFs can be found in Fig. S12D.

Fig. 6. Endogenous ECT1 has negligible ECT2/3/4-like function, despite similar expression pattern and subcellular localization. (*A-D*) Expression pattern of *ECT1p:gECT1-TFP* (gDNA) in aerial organs (*A*), main root (*B*), lateral root primordia (*C*) and emerging lateral roots (*D*) of 10-day-old seedlings. The expression mimics the pattern observed for ECT2/3/4 (Arribas-Hernández et al., 2020). (*E-F*) Intracellular localization of ECT1-TFP in meristematic cells of root tips in unchallegended conditions. Although the cytoplasmic signal is largely homogenous (*E*), sporadic foci (*F*) are frequently observed. Fig. S15A shows the integrity of the fluorescently tagged protein in independent transgenic lines assessed by protein blot. (*G*) Schematic representation of the *ECT1* locus. Exons are represented as boxes and introns as lines. The positions and identifiers of the T-DNA insertions assigned to the *ect1-1*, *ect1-2* and *ect1-3* alleles are marked, and the location of qPCR amplicons and hybridization probes (P) for analyses is indicated below. (*H*) Expression analysis of *ECT1* mRNA in wild type and T-DNA insertion lines by qPCR. Northern blot using the probe (P) marked in *G* detects *ECT1-TFP* mRNA, but not the endogenous *ECT1* transcript (Fig. S15B). (*I*) Morphological appearance of seedlings with or without *ECT1* in the different backgrounds indicated. Alternative allele combinations are shown in Fig. S15D. DAG, days after germination. Scale bars are: 1 cm in A, 1 mm in B, 100 µm in C, 10 µm in D-F, 1 mm in upper panels of I, and 1 cm in the lower panels.

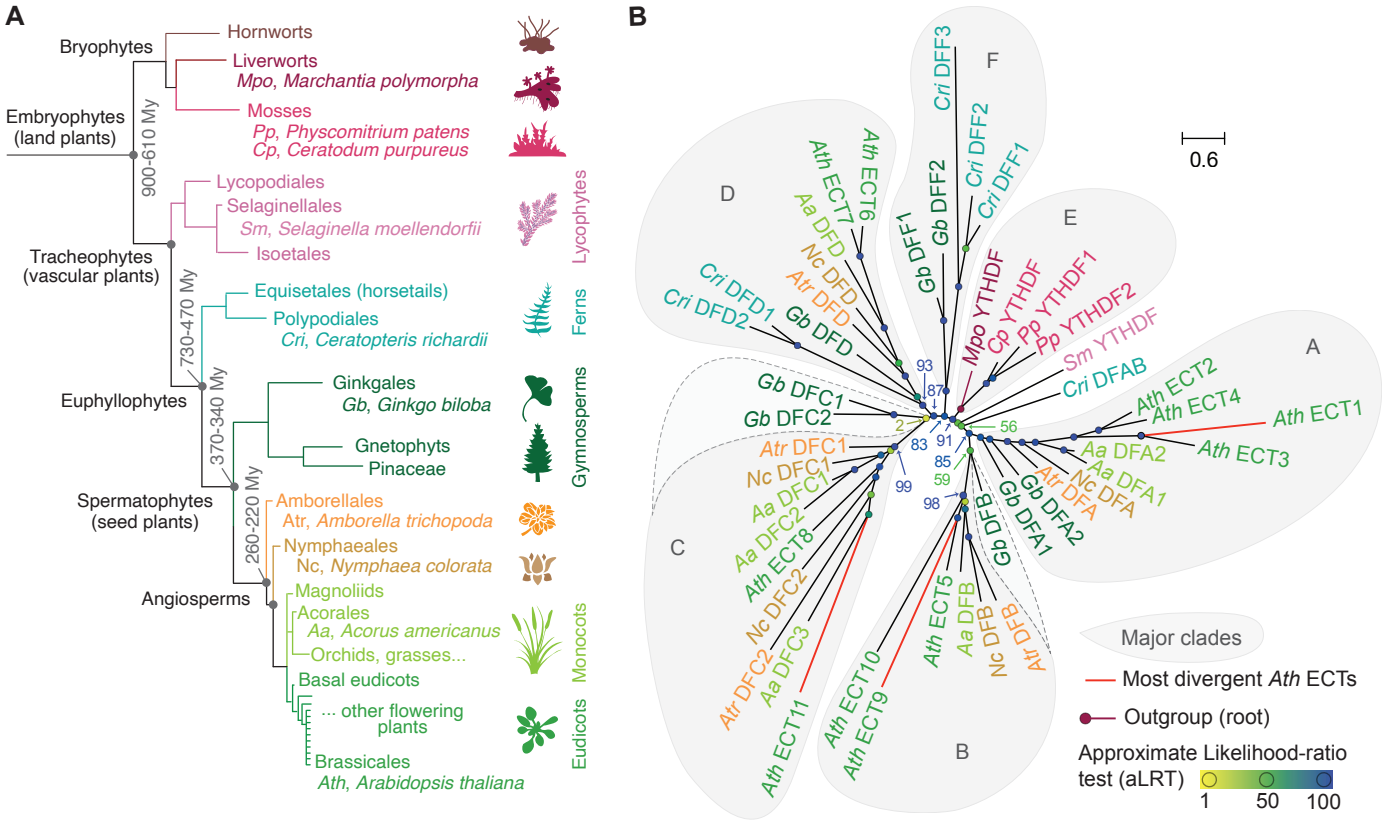

Fig. 1. Phylogenetic analysis of YTHDF proteins in land plants. (*A*) Schematic representation of land plant evolution and relative position of the species used in this study. The architecture of the diagram and the age (million years, My) indicated on some nodes are a simplified version of the tree from Su et al. (2021). The length of the branches has been adjusted for illustrative purposes. (*B*) Phylogenetic tree of YTHDF proteins in land plants based on the sequence alignment of regions with homology to the structured regions of the metazoan orthologs (Xu et al., 2015) (Datasets S1, S2). *Marchantia polymorpha* (*Mpo)* YTHDF is chosen as outgroup. *Arabidopsis thaliana* (*Ath*) YTHDF proteins are named after the original nomenclature for proteins containing an Evolutionarily Conserved C-Terminal Region (ECT) established by Ok et al. (2005). Proteins from other plant species adhere to the nomenclature established by Scutenaire et al. (2018), with small variations reflecting additional clades (DF-E, -F) and the split of the original DF-C clade into -C and -D. Statistical support is calculated as approximate likelihood-ratio test (aLRT) and indicated by colour-coded spheres on all nodes, but the most relevant values are specified. The length of the branches represents the evolutionary distance in number of amino acid substitutions per site relative to the scale bar. Color-coding and abbreviations of species as in *A*. Dotted lines highlight weakly supported affiliations to clades.

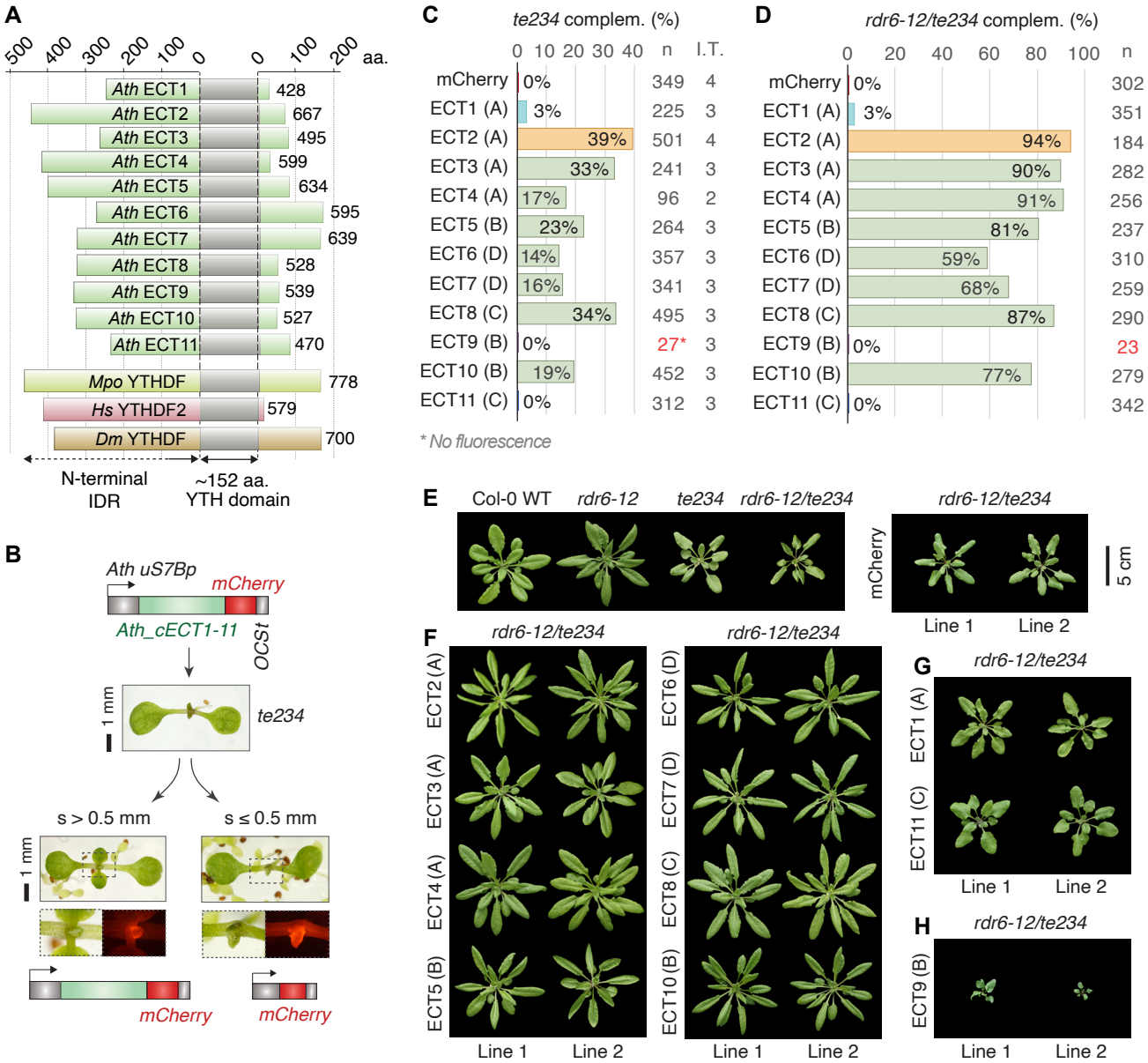

Fig. 2. Most but not all Arabidopsis YTHDF proteins can replace ECT2/3/4-function. (*A*) Diagram showing the relative length of the N-terminal IDRs of *Arabidopsis thaliana* (*Ath*) YTHDF proteins (ECT1-ECT11) together with *Marchantia polymorpha* (*Mpo*) YTHDF, *Homo sapiens* (*Hs*) YTHDF2, and *Drosphila melanogaster* (*Dm*) YTHDF. Numbers indicate the length of the proteins in amino acids (aa). (*B*) Strategy followed for the functional assay. *uS7Bp:cECT(X)-mCherry* constructs are introduced in *ect2-1/ect3-1/ect4-2* (*te234*) plants, and complementation rates are estimated by the percentage of primary transformants (T1) whose first true leaves have a size (s) exceeding 0.5 mm after 10 days of growth. The construct *uS7Bp:mCherry* is used as negative control. Examples of T1 seedlings expressing control and ECT2 constructs are shown. Dashed outlines are magnified below to show mCherry fluorescence in emerging leaves. (*C*) Weighed averages of the complementation percentages observed for each *uS7Bp:cECT(X)-mCherry* construct in 2-5 independent transformations (I.T.). Letters between parenthesis indicate the DF- clade assigned to each protein. Coloring highlights ECT2 (reference) and proteins with low or negligible complementation capacity (ECT1/9/11) for fast identification along the manuscript. Low transformation efficiency (T.E.) for ECT9 can be infered by the low number of transformants (n) highlighted in red, but detailed T.E. and raw complementation rates in each transformation can be found in Figs. S5 and S3*A* respectively. Additional data regarding this assay is shown in Fig. S4. (*D*) Same as *C* on a single transformation using the *rdr6-12/te234* background. (*E-H*) Rosette phenotype of 32-day old T1s bearing *uS7Bp:cECT(X)-mCherry* constructs in the *rdr6-12/te234* background, sorted according to their degree of complementation in three panels: *F* (high), *G* (residual) and *H* (non complementation and enhancement). Different genetic backgrounds and control transformants (mCherry) are shown in *E* as a reference. The scalebar (5 cm) applies to all plants in the four panels.

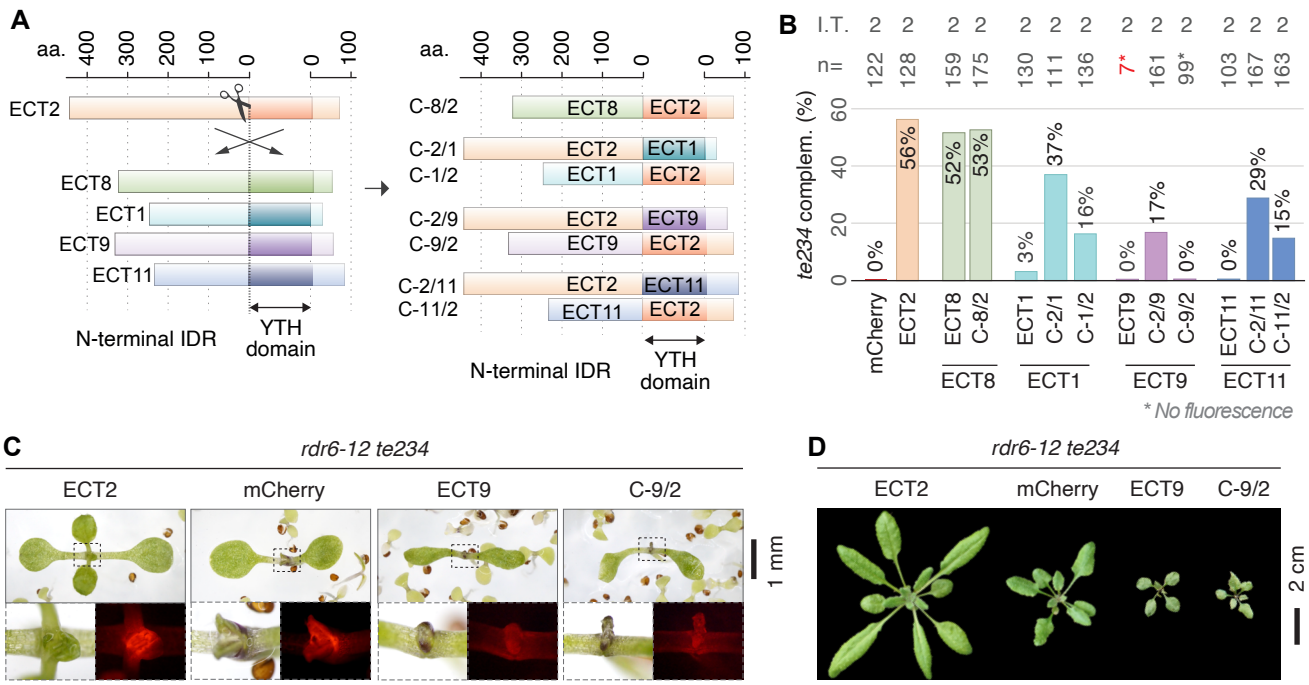

Fig. 3. Dissection of ECT1/9/11 functional regions. (*A*) Schematic representation of the strategy followed to express chimeras with N-terminal IDR and YTH domains of different ECT proteins. (*B*) Weighed averages of the complementation rates observed for each chimeric construct measured as in Fig. 2*B,C*. The raw complementation rates and transformation efficiencies of each independent transformation (I.T.) can be found in Figs. S3*B* and S5 respectively, and photographs of the scored 10-day-old T1 seedlings with mCherry fluorescence are shown in Fig. S8. (*C*) 10-day-old primary transformants of *rdr6-12/te234* with the indicated transgenes. Dashed outlines are magnified below to show mCherry fluorescencece. (*D*) Same genotypes as in *C* at 26 days after germination. Additional independent transgenic lines, developmental stages and controls for *C* and *D* can be found in Fig. S6. All transgenes are expressed from the *uS7B* promoter and contain mCherry fused to the C-terminus.

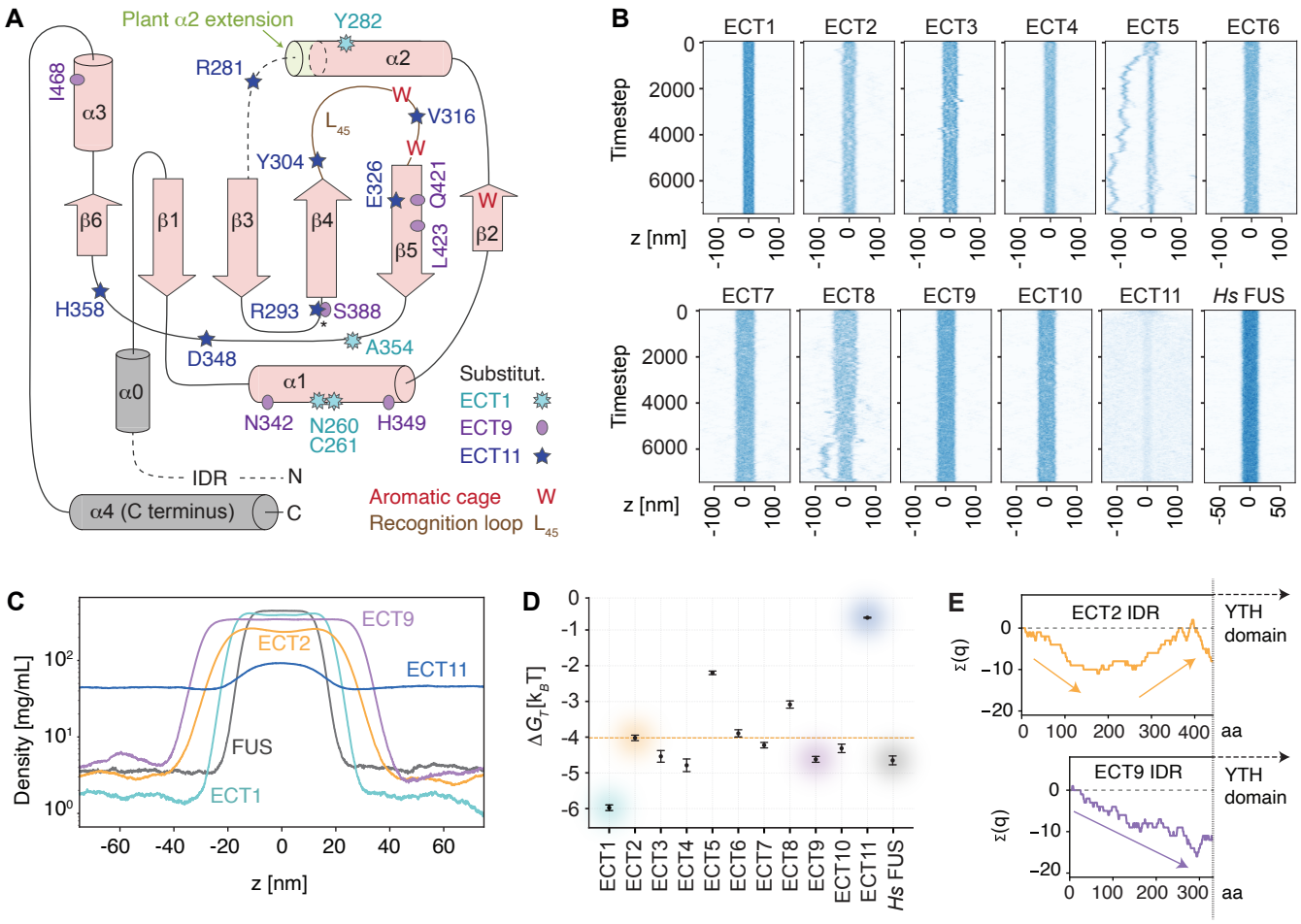

Fig. 4. Analysis of ECT1/9/11 functional regions. (*A*) Topology of the YTH domain of YTHDF1 with secondary structure elements drawn in pink (Xu et al., 2015). The two grey helices represent the only structural elements outside the YTH domain (as defined by Stoilov et al. (2002)): α0 immediately upstream, and α4 at the C-terminus of the protein. Superimposed on the diagrams are the amino acids in ECT1, ECT9 and ECT11 that constitute non-conservative substitutions compared to the majority of ECTs (see alignment in Fig. S9). The localization of these amino acids in the predicted 3D structures, predominantly on the surface, and the spatial integrity of the aromatic cage in the homology models of ECT1/9/11 can be found in Fig. S10. (*B*) Slab simulations of the IDRs of the indicated proteins showing their tendency to remain in condensed phase or disperse in solution. (*C*) Time-averaged density profiles for a selected group of the slab simulations shown in *B*. The remaining proteins are shown in Fig. S12*A*. (*D*) Calculated free energy change for the transition between aqueous solution and condensed phase (ΔG_{trans} = RT ln (c_{dilute} / c_{dense})) for the IDRs of the indicated proteins. (*E*) Cumulative charge distribution of the IDRs of the indicated proteins. The length in amino acids (aa) is at scale. The same analysis for all *Ath* ECTs can be found in Fig. S12*D*.

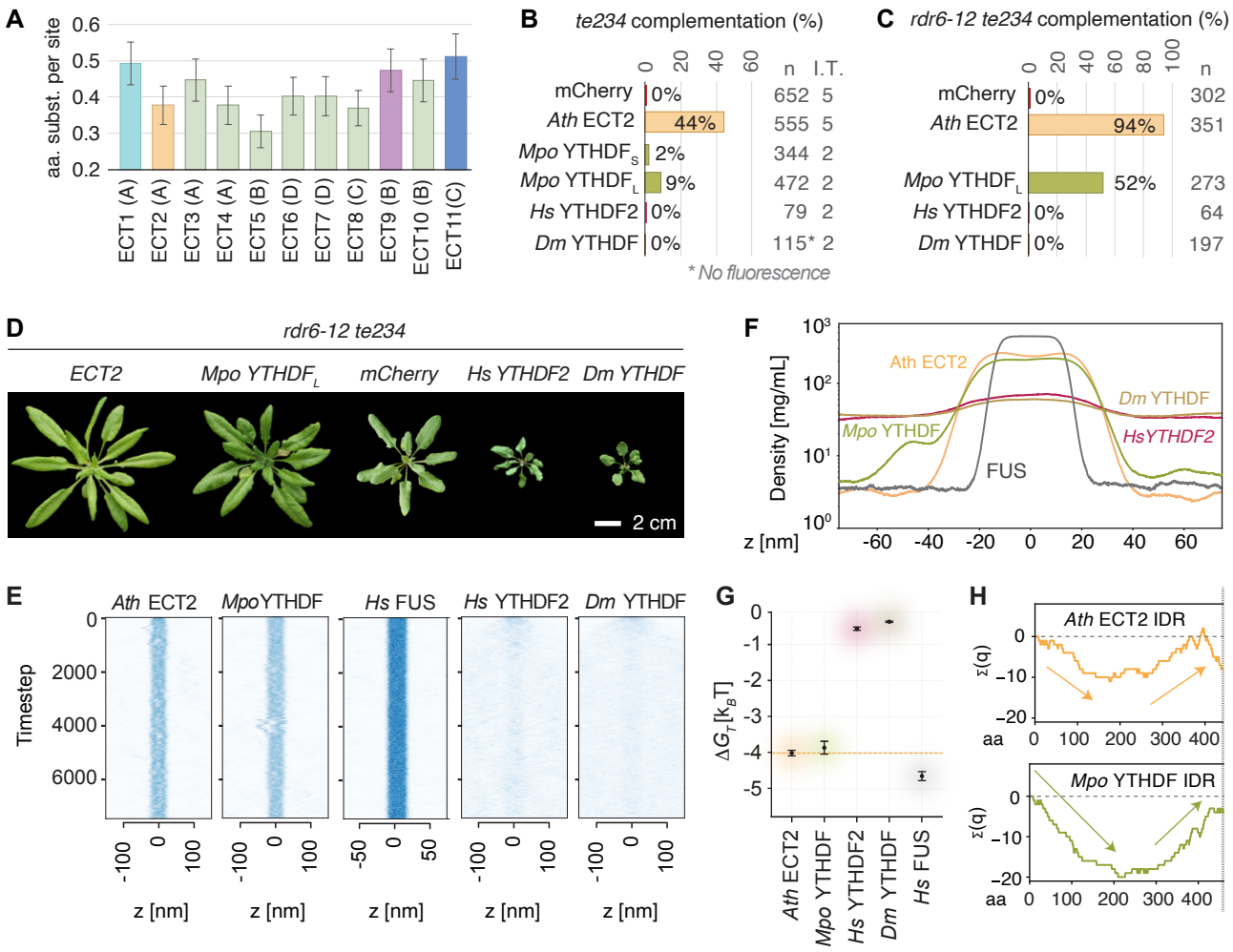

Fig. 5. Heterologous expression of *Marchantia polymorpha* (*Mpo*) YTHDF in Arabidopsis rescues loss of ECT2/3/4 function, while *Homo sapiens* (*Hs*) or *Drosophila melanogaster* (*Dm*) YTHDF proteins aggravates it. (*A*) Evolutionary divergence among ECT proteins estimated by pairwise comparison of the proteins analyzed in Fig. 1*B*. (*B*) Weighed averages of the complementation rates observed for each of the indicated constructs (cDNA fused to mCherry under the control of the *uS7B* promoter) as in Fig. 2*C* in the *te234* background. S (short) and L (long) refer to the two different YTHDF spliceforms found among *M. polymorpha* cDNAs. Transformation efficiencies and raw complementation rates in each transformation can be found in Figs. S5 and S3 respectively. (*C*) Same as in *B* but from one transformation in the *rdr6-12/te234* background. (*D*) 32-day-old primary transformants of *rdr6-12/te234* expressing the heterologous transgenes. Additional controls and developmental stages are shown in Fig. S6. (*E*) Slab simulations of the IDRs of the indicated proteins as in Fig. 4*B*. (*F*) Time-averaged density profiles for the slab simulations in *E*. (*G*) Calculated free energy change for the IDRs of the indicated proteins as in Fig. 4*D*. (*H*) Cumulative charge distribution of the IDRs of the indicated proteins as in Fig. 4*E*. The same analysis for all arabidopsis ECTs and animal YTHDFs can be found in Fig. S12*D*.

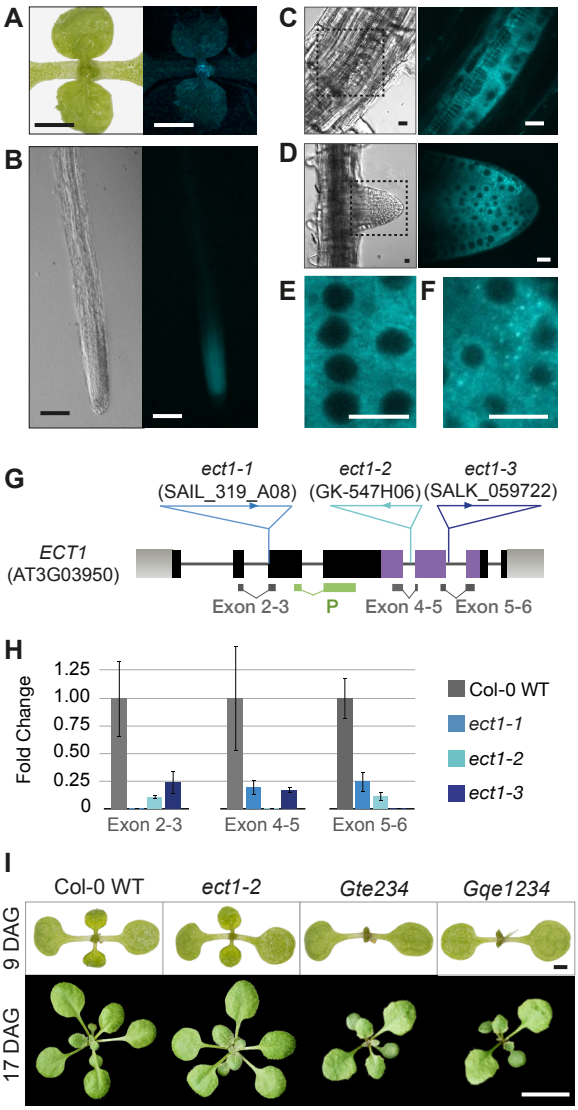

Fig. 6. Endogenous ECT1 has negligible ECT2/3/4-like function, despite similar expression pattern and subcellular localization. (*A-D*) Expression pattern of *ECT1p:gECT1-TFP* (gDNA) in aerial organs (*A*), main root (*B*), lateral root primordia (*C*) and emerging lateral roots (*D*) of 10-day-old seedlings. The expression mimicks the pattern observed for ECT2/3/4 (Arribas Hernández et al., 2020). (*E-F*) Intracellular localization of ECT1-TFP in meristematic cells of root tips in unchallegended conditions. Although the cytoplasmic signal is largely homogenous (*E*), sporadic foci (*F*) are frequently observed. Fig. S15*A* shows the integrity of the fluorescently tagged protein in independent transgenic lines assessed by protein blot. (*G*) Schematic representation of the *Ath_ECT1* locus. Exons are represented as boxes and introns as lines. The positions and identifiers of the T-DNA insertions assigned to the *ect1-1*, *ect1-2* and *ect1-2* alleles are marked, and the location of qPCR amplicons and hybridization probes (P) for analyses is indicated below. (*H*) Expression analysis of *ECT1* mRNA in wild type and T-DNA insertion lines by qPCR. Northern blot using the probe (P) marked in *G* detects *ECT1-TFP* mRNA, but not the endogenous *ECT1* transcript (Fig. S15*B*). (*I*) Morphological appearance of seedlings with or without *ECT1* in the different backgrounds indicated. Alternative allele combinations are shown in Fig. S15*D*. DAG, days after germination. Scale bars are: 1 cm in *A*, 1 mm in *B*, 100 μm in *C*, 10 μm in *D-F*, 1 mm in upper panels of *I*, and 1 cm in the lower panels.