LARP6 proteins in plants

Cécile Bousquet-Antonelli

MINI REVIEW

LARP6 PROTEINS IN PLANTS

Cécile BOUSQUET-ANTONELLI[1,2,*]

[1]CNRS LGDP-UMR5096, 58 Av. Paul Alduy 66860 Perpignan, France

[2]Université de Perpignan Via Domitia, LGDP-UMR5096, 58 Av. Paul Alduy 66860 Perpignan France

Correspondance : cecile.antonelli@univ-perp.fr

Abstract

RNA binding proteins, through control of mRNA fate and expression, are key players of organism development. The LARP family of RBPs sharing the La motif, are largely present in eukaryotes. They classify into five subfamilies which members acquired specific additional domains, including the RRM1 moiety which teams up with the La motif to form a versatile RNA binding unit. The LARP6 subfamily has had a peculiar history during plant evolution. While containing a single LARP6 in algae and non-vascular plants, they expanded and neofunctionalized into three subclusters in vascular plants. Studies from *Arabidopsis thaliana*, support that they acquired specific RNA binding properties and physiological roles. In particular LARP6C participates, through spatiotemporal control of translation, to male fertilization, a role seemingly conserved in maize. Interestingly, human LARP6 also acts in translation control and mRNA transport and similarly to LARP6C which is required for pollen tube guided elongation, is necessary to cell migration, through protrusion extension. This opens the possibility that some cellular and molecular functions of LARP6 were retained across eukaryote evolution. With their peculiar evolutionary history, plants provide a unique opportunity to uncover how La-module RNA binding properties evolved and identify species specific and basal roles of the LARP6 function. Deciphering of how LARP6, in particular LARP6C, acts at the molecular level, will foster novel knowledge on translation regulation and dynamics in changing cellular contexts. Considering the seemingly conserved function of LARP6C in male reproduction, it should fuel studies aimed at deriving crop species with improved seed yields.

Introduction

RNA Binding Proteins (RBPs) are key players of posttranscriptional regulations, involved in the control of mRNA fate, they intervene at all steps of an mRNA life, influencing their coding sequence, life duration, localization and translation. As mRNA regulators, RBPs provide a

reactive and fine mean to orchestrate and modify the pattern of gene expression to answer the cellular needs in response to developmental or environmental cues. RBPs recognize through their RNA Binding Domain(s) (RBDs), RNA regulatory motifs that are primary or secondary structure elements or chemical modifications and form with mRNAs, ribonucleoprotein particles (mRNPs). The RNA Recognition Motif (RRM) is the most represented and dominant classical RBD in eukaryotes, including plants. But a variety of other RDBs exist among which the La Motif (LaM), an eukaryotic domain that adopts a double winged helix fold [1,2]. Firstly identified in humans, on the genuine La protein (or La autoantigen) [3], the LaM is an ancestral domain that emerged early after the *archea-eukarya* radiation, and is found on hundreds of factors, that form the superfamily of LA and Related Proteins (LARPs) [4]. The LaM is often tethered by a short linker to a downstream RRM (RRM1) with which it allies to form the La- module, a single RNA binding unit. During *eukarya* evolution, the LARP family amplified and evolved into subclusters which members acquired additional group-specific domains that can bind RNA, team up with the La-module or mediate protein-protein interactions. In the past 10 years, LARPs attracted much attention and functional studies of members of each subgroup: genuine La, LARP1, 4, 6 and 7, are being conducted in mammals, invertebrates, fungi or plants. LARPs endorse many crucial roles in organism growth and development, as well as response to stress [5] and in mammals, several of them are recognized as cancer-associated factors [6]. The genuine La and LARP7 proteins target diverse types of polymerase III encoded RNAs to promote their maturation, folding and assembly into RNPs. Other LARPs (LARP1, 4 and 6) as well as the cytoplasmic isoform of the genuine La, bind mRNAs to control their localization, stability and/or translation.

In fibroblasts, human LARP6 (HsLARP6) coordinates the translation of type-I collagen subunits at endoplasmic reticulum [7]. Through direct binding to a 5’ stem-loop (5’-SL) shared by the 5’-untranslated regions (5’-UTRs) of the α1 and α2 transcripts from type-I collagen [8,9], HsLARP6 docks its targets to vimentin or non-muscle myosin filaments hence participating in their transport. It also acts either as repressor or activator of collagen translation through recruitment of distinct effectors to its target transcripts [9–12]. In migratory mammalian cells, HsLARP6 also transports ribosomal protein mRNAs from cell body to protrusion where their translation is locally increased, leading to increase ribosome biogenesis and protein synthesis [13]. HsLARP6 was found to be up regulated in aggressive carcinomas, and proposed to participate in cancer progression [13]. In zebrafish (*Danio rerio*), that contains two paralogs, LARP6 proteins are required for normal oocyte development, chorion formation and egg activation, but the molecular bases of this are so far unknown [14].

LARP6 proteins appear as fascinating multifaceted factors, which roles in translational control and mRNA transport, likely fuels diverse physiological processes. Plant proteins

seem to be no exception to the rule, at least based on a yet too scarce knowledge of their evolutionary history, molecular and physiological roles that I will review here.

A plant-specific evolutionary history

LaM proteins are present in all eukaryotes (except protista from the *Plasmodium* genus) but not all subfamilies of LARPs are represented in all of them. LARP6 proteins are only present in three lineages: Stramenopiles, Metazoan and Viridiplantae [4,15,16]. Embryophytes ("land plants") and Chlorophytes ("green algae") that form the Viridiplantae ("green plants") clade, lack LARP4 and LARP7 subgroups while their LARP6 family expanded (Figure 1A).

All LARP6 proteins carry a La-module, with a subfamily specific RRM1, tagged as RRM-L3b on vertebrate proteins, and RRM-L3a on other proteins including plant ones (Figure 1B, 1C). Primary sequence comparison of plant LARP6 La-modules, unraveled their specific evolutionary history. In non-vascular plants (green algae and mosses (bryophytes)), LARP6s are generally encoded by a single gene while in vascular plants, the family expanded and contains three to six members, with most dicotyledone species having three and monocotyledone six. Vascular plant proteins evolved into distinct subclusters that were labeled 6A, 6B and 6C (Figure 1A). Every vascular plant carries at least one member of each subgroup, with in most cases, a single LARP6A and one to three 6B and 6C proteins. Notably, members of the LARP6B and 6C subgroups acquired a PAM2 (PABP interacting motif 2), a structure peptide mediating direct interaction with the poly(A) binding protein (PABP) [17]. This acquisition seems to be coupled with a significant reorganization of their La-module, that have a divergent LaM, that contrarily to that of non-vascular plant LARP6 and LARP6A, underwent non conservative changes on the six amino acids known in genuine La and other LARPs, to participate in RNA interaction [1,2,16]. Moreover, while RRM-L3a are predicted to have an extended loop 3 of invariant length, plant LARP6B and 6C appear to carry loop 3 of more variable lengths (Figure 1B). This observation, based on *in silico* structure predictions should be considered with caution, until confirmed through biophysical approaches, but nonetheless, together with sequence divergence of the LaM, points to possible subgroup specific RNA-binding properties [4]. Vertebrate LARP4 proteins form the second group of LARPs that gained a PAM2 across evolution. This acquisition also associates with a reorganization of their La-module, that was found experimentally on human LARP4 (HsLARP4), to significantly alter their RNA binding properties. Strikingly this La-module only modestly contributes to RNA binding and assumes a non-canonical linear fold, conversely to other LARP La-modules where LaM and RRM1 form a V-shaped RNA binding unit [5,9,18].

Biochemical properties of plant LARP6 conserved domains

Plant LARP6 are, as other LARP6s, modular proteins, sharing together with their La-modules, the LSA (LaM and S1 Associated) domain located at their C-terminus. Viridiplantae LARP6 additionally acquired RG repeats placed between La-module and LSA and vascular plant proteins, a PAM2 located at their N-terminus [4,15,16] (Figure 2).

The La-module is a highly versatile domain that employs various molecular strategies to recognize and bind its RNA targets [19]. Rather than a bipartite, it can be viewed as a tripartite RNA binding unit composed of the LaM, RRM1 and linker region which also plays a crucial role. The LaM and RRM1 moieties adopt their globular fold independently from each other and, depending on the linker size and orientation, differently position one relative to the other [5,19]. Biophysical studies on plant LARP6 La-module are still awaited but HsLARP6 La-module was explored [9,20]. Globally its LaM and RRM1 (/RRM-L3b) adopt canonical tertiary folds, but with specific structural variations [9]. The most striking variation of the LaM is that of wing 2 which terminates the structure (Figure 1C). As compared to genuine La, which solution structure is well characterized, HsLARP6 wing 2 is longer and consequently, the interdomain linker between LaM and RRM1 shorter (2 amino acids versus 11 in HsLa). Also wing 2 orientation points to a different direction, likely positioning the RRM1 differentially with regard to the LaM. This has a significant impact on RNA binding, as when its linker was replaced with that of human La, HsLARP6 La-module lost its ability to bind its target RNA. RRM1 which carries two additional helixes of yet unknown importance (Figure 1C), is absolutely necessary for RNA binding, as here again a chimeric La-module with a genuine La RRM1, no longer associates with RNA [9]. Whether the La-modules of plant LARP6 also display these structural specificities, in particular whether they carry this peculiar wing 2 and 2-aa long linker, is difficult to tell based on sequence alignments. Anyhow, divergences exist. Although able to bind RNA [16], the La-modules from *Arabidopsis thaliana* (At) LARP6A and C proteins are not able to bind the 48-nt long 5'-SL of collagen mRNAs [9]. In addition a chimeric La-module formed with HsLARP6 LaM and AtLARP6C RRM1 (/RRM-L3a), no longer associates to the 5'SL RNA [19].

In vitro AtLARP6C La-module shows a strong specificity towards oligo(U) homopolymers and binds with more affinity probes with uridine stretches interspersed with cytidines [16,21]. RNA-recognition by 6C La-module depends on the length of the probe (with a 20-nt probe being a better substrate than a 10-nt probe) but binds RNA regardless of their 3' extremity (3'-OH or 3'-PO4) which supports that it could interact with terminal or internal sequences on RNA *in vivo. In planta* functional analyses, corroborated these findings, as mRNA substrates of AtLARP6C were found to share in their 5'-UTRs a (UUC/UUUC) box that AtLARP6C preferentially binds *in vitro* [21].

AtLARP6A La-module shows preference towards oligo(A) probes regardless of their length or 3'-end nature, but seems less specific than that of AtLARP6C, since it interacts also with

oligo(U) and oligo(C) or (G) homopolymers albeit with lower affinity [16]. So far, an exploration of RNA binding preferences of a plant LARP6B is lacking, as in striking contrast with 6A or 6C, the La-module of the Arabidopsis protein failed to be produced and purified from *Escherichia coli*.

In addition to the family specific La-module, the LSA is also a hallmark of LARP6 proteins. Always located at the very C-terminal end, it is a 20 to 30 amino acids long motif highly conserved across evolution. On animal proteins the LSA portion constitutes the C-terminal part of a larger domain, called SUZ-C (SUZUKI-C) shared by other proteins, including those from the CSP1 (cold-shock response protein 1) [4] and SZY-20 families [22] (Figure 2A). LARP6 proteins from the green lineage only carry the LSA portion of SUZ-C and systematically have the 3 amino acid "PRM" insertion as compared to other eukaryotes (Figure 2B, 2C). No other plant protein seems so far to carry a LSA or SUZ-C domain.

The role of the LSA from plant proteins is yet unknown and in animal, its roles are still unclear. SUZ-C/LSA domains are often found on proteins carrying RBDs, such as S1-like on CSP1 proteins, and on SZY-20, it was proposed to mediate RNA interaction [22]. However, its deletion on HsLARP6 does not seem to impede binding to the collagen 5'-stem-loop [8]. Actually, it was rather proposed to mediate protein-protein interactions. The C-terminal domain of HsLARP6 starting immediately downstream to the La-module and encompassing the SUZ-C region, is necessary and sufficient to co-precipitate HsLARP6 protein partners from crude extracts, what is no longer the case when the LSA part is delete [10,12]. But although indicative, this experimental approach does not permit to conclude on a role for the LSA/SUZ-C domain in directly mediating protein-protein interaction. Further work is hence needed to uncover the biochemical properties of this conserved domain.

PABP is a universal eukaryotic RBP, that coats the poly(A) tail of mRNAs and is a crucial enhancer of translation and a regulator of poly(A) tail length and mRNA stability. The C-terminal MLLE domain of PABP, is a conserved platform where PAM2-carrying factors associate. PAM2 which consensus is “xxLxxxAxx(F/W)xP”, is found on dozens of proteins, including members of the vertebrate LARP4 subgroup [16,17]. *In vitro* and *in vivo* studies demonstrated that PAM2 motifs of AtLARP6B and 6C are necessary and sufficient to mediate a direct interaction with the MLLE domain of plant PABP [16]. Interestingly, the PAM2 from vertebrate LARP4, mediates MLLE interaction but is also necessary although not sufficient to bind RNA [18,23]. Strikingly, position 10 is crucial for RNA binding and at least on HsLARP4, must be a tryptophan and not a phenylalanine residue [18]. Whether PAM2 motifs from plant LARP6B and C are also RNA binding surfaces was not tested, but their position 10 is always a phenylalanine and not a tryptophan residue, suggesting that it might not be the case.

Recently a deeper evolutionary analysis uncovered the existence of RG repeats located upstream the LSA of embryophyte LARP6, including LARP6 proteins from mosses and A, B and C type proteins from vascular plants [15] (Figure 2B, C). RGG/RG repeats are widespread on eukaryotic proteins and often found on factors involved in mRNA regulations. Their most widely reported biochemical property, is their ability to bind RNA, but they can also support protein – protein interactions or be implicated in the regulation of protein localization [24]. An interesting feature of RGG/RG repeats is that their arginine residues are preferred sites for methylation, a posttranslational modification found to regulate properties, activities and functions of the repeats. Plant LARP6 proteins seem to be versatile factors, which molecular and physiological functions are altered in response to cellular or environmental cues (see below), it is tempting to speculate that arginine methylation could participate in their functional regulation.

Physiological and molecular functions of plant LARP6 proteins

To date functional data are available on vascular plant LARP6 proteins, mostly from Arabidopsis and *Zea maize* [16,21,25]. Arabidopsis contains three LARP6 proteins (one of each type) [16] and maize six, with one 6A, three 6B and two 6C [25] (Figure 2B). In Arabidopsis, LARP6 genes show specific expression profiles, with in particular *LARP6C* exclusively expressed in pollen from the onset of gametogenesis to the end of fertilization process [21,26,27]. A striking feature of AtLARP6C is that it is a multifaceted protein likely to endorse distinct molecular roles along male gametogenesis and progamic phase. During gametogenesis, AtLARP6C moves for a nucleolar in uninucleate microspore to a cytosolic localization in the vegetative cell of bi- (BCP) and tricellular pollen (TCP). During the very last step of gametogenesis between TCP and MPG (mature pollen grain), AtLARP6C transitions from the soluble fraction of the cytosol to aggregates, that also contain PABP and were proposed to be mRNP granules. Moreover, while excluded from sperm cells, it labels the endo-plasma membrane that engulfs the sperm cells (SCs) (including the cytoplasmic connection bridging SC and VCN) [28] and coats the outside of vegetative cell nucleus (VCN). Although modestly, AtLARP6C is required for proper gametogenesis, but its cellular and molecular functions in this process yet remain unexplored. Its contribution to the fertilization process from grain hydration to sperm cell delivery was explored in more details. During the early steps of the progamic phase, *atlarp6c* loss-of-function mutants display altered germination dynamics and pollen tube elongation rates. Following penetration in the pistil, *larp6c*-deficient pollen tubes although elongating at wild-type pace, fail to target ovules at normal frequency: they display a clear, although not fully penetrant, guidance defect. When they reach and penetrate the ovule, pollen tubes are seemingly able to burst and deliver sperm cells without any contribution of AtLARP6C [21].

The involvement of AtLARP6C in the fertilization process is fueled by its function as RNA binding protein and regulator of mRNA fate in the cytoplasm. Its client mRNAs in MPG code for functions known to be necessary for germination and polarized cell growth that is at the basis of the PT guidance process. In particular AtLARP6C binds actors of fatty acid and glycerolipid synthesis and homeostasis and actors of vesicular trafficking, two cellular processes that are significantly perturbed in the absence of AtLARP6C during pollen tube guided growth. Experiments conducted with one of AtLARP6C target transcripts as model (*MGD2*: monogalactosyldiacylglycerol synthase 2), support that through direct binding to a (UUC/UUUC) 20nt-long motif at 5'-UTR, it controls their translational status: endorsing an inhibitor role in MPG and shifting to an activator at hydration and during the fertilization process. Interestingly in pollen tubes, AtLARP6C is present both in the cytosol and in cytoplasmic foci, proposed to be mRNP granules, and is likely to bind microtubule cytoskeleton. This dual localization opens the interesting possibility that AtLARP6C is both activator and inhibitor of translation in growing PTs. The authors propose that pollen tube anisotropic growth and guidance could, as it is the case for axons in neurons [29], rely on localized translation, where silent mRNAs must be transported to translational hotspots where their protein is needed to answer cellular needs and propose that AtLARP6C could be one actor of this process [21].

In maize, ZmLARP6C1, the pollen specific LARP6, is also required during the progamic phase, with loss-of-function mutants showing reduced male transmission that is the consequence of less competitive pollen. Indeed *zmlarp6c1*-deficient pollens show altered germination dynamics and PT elongation rates [25]. An assessment of the fertilization ability of mutant maize pollens was not conducted, but already the similarities with Arabidopsis phenotypes during the early steps of the progamic phase, suggest that LARP6C molecular and cellular roles could be conserved in maize and more widely in angiosperms (flowering plants).

AtLARP6A, although ubiquitously expressed, shows its highest mRNA accumulation in MPG where it is more expressed than *AtLARP6C* and is in the top 300 to 500 pollen genes depending on transcriptomic data [26,27,30]. AtLARP6A seems unlikely to work redundantly with AtLARP6C. Not only loss of 6A does not appear to alter *atlarp6c* phenotypes, but also the proteins show only partial overlapping distribution in MPG and pollen tubes [21]. In MPG, AtLARP6A accumulates in the vegetative cell cytoplasm, but does not seem to form foci [21], what is consistent with the observation that it does not relocates in mRNP aggregates in onion epidermis cells following stress exposure [16]. It also accumulates in the cytoplasm of sperm cells and seemingly in the VCN of MPG. In pollen tubes, its distribution does not change, except in VCN where it seems to accumulate in nucleoli, a distribution that would need further exploration but which is consistent with the localization of AtLARP6A in onion

epidermis cells [16,21]. The role of AtLARP6A in pollen biology and fertilization is yet to explore, but *atlarp6a*-deleted pollens seem to be transmitted at wild-type frequency what is surprising considering its high expression levels in pollen.

Conversely to the other two Arabidopsis LARP6, *AtLARP6B* mRNA has a ubiquitous expression profile across plant development, but is not detected in pollen nor in ovules [21]. AtLARP6B, is a direct interactant of PABP, and likely an RNA binding protein. In onion epidermis cells, it accumulates in the nucleolus, while excluded from the nucleoplasm and upon hypoxia, it relocates to cytoplasmic aggregates, also containing PABP and most likely to be mRNP granules [16]. AtLARP6B was hence proposed to have function in mRNA metabolism, at least under certain environmental or cellular conditions.

The situation is less obvious for AtLARP6A. It does not carry a PAM2 and does not appear to belong to mRNP aggregates, hence it is so far impossible to tell whether it acts to regulate mRNAs or other type of transcripts, such as non-coding RNAs.

Conclusion

LARP6 proteins have had a peculiar history across plant evolution. While basal and non-vascular plants have a single "canonical" LARP6, vascular plant proteins expanded and evoled into subcusters. Did and why vascular plants acquire novel and multiple LARP6 functions? Out of its expression profile, we can tell that AtLARP6B is not involved in male reproduction and there is no aggravation of *atlarp6c*-pollen phenotypes due to the loss of LARP6A. This supports that 6C fulfills physiological functions distinct to that of 6B and 6A. Nonetheless, although animal LARP6 has no PAM2 and a canonical La-module, one can find similarities between their cellular roles. HsLARP6 as likely AtLARP6C, is involved in the transport and translation status control of its bound mRNAs. Their role in mRNA transport could rely on their interaction with cytoskeleton, as both Arabidopsis and Human LARP6 form cytoplasmic aggregates closely tracking the microtubule filaments. And HsLARP6 is a positive regulator of cell protrusion elongation, a cellular function comparable to that of AtLARP6C in pollen tube elongation and guided growth [13,21]. It is also tempting to speculate that such cellular role for LARP6 is conserved in zebrafish. Indeed, Hau and collaborators propose that through regulation of mRNA metabolism, LARP6 proteins are necessary for the formation of microvilli that are cellular extensions protruding from the developing oocytes and their surrounding granulosa cells [14]. These observations suggest that although adapted to specific cellular/species contexts, these functions were retained in eukaryote evolution and might represent some basal LARP6 functions.

Hence do LARP6A and/or LARP6B have plant-specific functions? A detailed characterization is necessary to answer this and understand why plant acquired several LARP6 during evolution. But study of the LARP6 function in algae and non-vascular plants, that have a

single protein, shall be also highly informative. We will learn, at a whole organism level, what are the physiological basal roles of this function and if it is essential for life. We should hint on the basal cellular and molecular roles of LARP6, in particular know if spatiotemporally controlling translation is one of its basic features. Also, the deciphering and comparison of the molecular modes of RNA binding from proteins from each group (LARP6, LARP6A, 6B and 6C), will give valuable insights on how La-modules evolved.

LARP6 proteins so far appear as translational switches adapting their inhibitor/activator rheostat in response to cellular cues at least during male reproduction. We now need to explore how, in particular focusing on dynamic posttranslational modification profiles and on the identification of the repertoires of LARP6 cofactors in different contexts. Such studies, in particular if conducted with AtLARP6C, should participate in the discovery of novel actors of male reproduction. Since LARP6C function appears conserved in maize [25], it should widen our understanding of crop reproduction and seed production.

Perspective

- LARP6 function was conserved across eukaryote evolution but plant proteins expanded, evolved and acquired novel protein domains supporting that they gained novel important molecular and physiological functions. Conservation of LARP6C function between Arabidopsis and maize supports a crucial role in plant reproduction.
- LARP6 proteins are recognized as an important family of mRNA regulators that sustain important physiological roles at least in plant reproduction and mammalian oncogenesis. LARP6 functions in the anisotropic growth of cells and as regulator of mRNA transport and translation can be viewed as basal conserved functions in eukaryotes.
- A characterization of LARP6 proteins from non-vascular plants could give precious insights into the evolution of their RNA binding modes and a deeper understanding of LARP6 basal functions. Deciphering how LARP6 acts as translational switch will participate in fostering novel knowledge on translation dynamics in response to the cellular context. A deeper characterization of AtLARP6C cofactors could foster knowledge on novel actors of plant reproduction including in crop species.

Competing interests: The author declares not competing interest.

Funding: Research on plant LARP6 in my team was funded by the CNRS and University of Perpignan Via Domitia (UPVD). UPVD doctoral school ED 305 awarded a phD grant, UPVD a BQR (Bonus Qualité Recherche) grant and CNRS a collaborative PICS (LARP&STRESS, grant no; 6170). Studies on LARP proteins are set within the framework of the “Laboratoires

d'Excellence (LABEX)" TULIP (ANR-10-LABX-41) and of the "Ecole Universitaire de Recherche" TULIP-GS (ANR-18-EURE-0019).

Acknowledgments: I would like to thank Jean-Marc Deragon (LGDP, Perpignan), Lian Zhou (Southwest University, Chongqing, China) and John E. Fowler (Oregon State University, Corvallis, USA) for their help to retrieve maize protein sequences and accession numbers. Thank you to Rémy Merret (LGDP, Perpignan) for critical reading of the manuscript.

Abbreviations:

mRNP, mRNA ribonucleoprotein particle; RBP, RNA binding protein; RBD, RNA binding domain; RRM, RNA recognition motif; LaM, La motif; LARP, La and related protein; 5'-UTR, 5' untranslated region; UNM, uninucleate microspore; BCP, bi-cellular pollen; TCP, tri-cellular pollen; MPG, mature pollen grain; PT, pollen tube.

[1] Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S, et al. (2004) Structural analysis of cooperative RNA binding by the La motif ans central RRM of human La protein. Nat. Struct. Mol. Biol. **11**, 323–329.

[2] Kotik-Kogan O, Valentine ER, Sanfelice D, Conte MR, Curry S. (2008) Structural analysis reveals conformational plasticity in th erecognition of RNA 3' ends by the human La protein. Structure. **16**, 852–862.

[3] Wolin SL, Cedervall T. (2002) The La Protein. Annu. Rev. Biochem. [Internet]. **71**, 375–403. Available from: http://www.annualreviews.org/doi/10.1146/annurev.biochem.71.090501.150003

[4] Bousquet-Antonelli C, Deragon J-M. (2009) A comprehensive analysis of the La-motif protein superfamily. RNA. **15**.

[5] Maraia RJ, Mattijssen S, Cruz-Gallardo I, Conte MR. (2017) The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. Wiley Interdiscip. Rev. RNA [Internet]. **8**. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28782243

[6] Sommer G, Heise T. (2021) Role of the RNA-binding protein La in cancer pathobiology. RNA Biol. [Internet]. **18**, 218–236. Available from: https://www.tandfonline.com/doi/full/10.1080/15476286.2020.1792677

[7] Zhang Y, Stefanovic B. (2016) LARP6 Meets Collagen mRNA: Specific Regulation of Type I Collagen Expression. Int. J. Mol. Sci. [Internet]. **17**, 419. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27011170

[8] Cai L, Fritz D, Stefanovic L, Stefanovic B. (2010) Binding of LARP6 to the conserved 5' stem-loop regulates translation of mRNAs encoding type I collagen. J. Mol. Biol.

[Internet]. **395**, 309–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19917293

[9] Martino L, Pennell S, Kelly G, Busi B, Brown P, Atkinson RA, et al. (2015) Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module. Nucleic Acids Res. [Internet]. **43**, 645–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25488812

[10] Cai L, Fritz D, Stefanovic L, Stefanovic B. (2010) Nonmuscle myosin-dependent synthesis of type I collagen. J. Mol. Biol. [Internet]. **401**, 564–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20603131

[11] Challa AA, Stefanovic B. (2011) A novel role of vimentin filaments: binding and stabilization of collagen mRNAs. Mol. Cell. Biol. [Internet]. **31**, 3773–89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21746880

[12] Vukmirovic M, Manojlovic Z, Stefanovic B. (2013) Serine-threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. Mol. Cell. Biol. [Internet]. **33**, 3893–906. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23918805

[13] Dermit M, Dodel M, Lee FCY, Azman MS, Schwenzer H, Jones JL, et al. (2020) Subcellular mRNA Localization Regulates Ribosome Biogenesis in Migrating Cells. Dev. Cell [Internet]. **55**, 298-313.e10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33171110

[14] Hau HTA, Ogundele O, Hibbert AH, Monfries CAL, Exelby K, Wood NJ, et al. (2020) Maternal Larp6 controls oocyte development, chorion formation and elevation. Development [Internet]. **147**. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32054660

[15] Deragon J-M. (2021) Distribution, organization an evolutionary history of La and LARPs in eukaryotes. RNA Biol. **18**, 159–167.

[16] Merret R, Martino L, Bousquet-Antonelli C, Fneich S, Descombin J, Billey É, et al. (2013) The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of la-related proteins. RNA. **19**.

[17] Xie J, Kozlov G, Gehring K. (2014) The “tale” of poly(A) binding protein: The MLLE domain and PAM2-containing proteins. Biochim. Biophys. Acta - Gene Regul. Mech. [Internet]. **1839**, 1062–1068. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1874939914002119

[18] Cruz-Gallardo I, Martino L, Kelly G, Atkinson RA, Trotta R, De Tito S, et al. (2019) LARP4A recognizes polyA RNA via a novel binding mechanism mediated by disordered regions and involving the PAM2w motif, revealing interplay between PABP,

LARP4A and mRNA. Nucleic Acids Res. [Internet]. **47**, 4272–4291. Available from: https://academic.oup.com/nar/article/47/8/4272/5366477

[19] Dock-Bregeon A-C, Lewis KA, Conte MR. (2021) The La-related proteins: structures and interactions of a versatile superfamily of RNA-binding proteins. RNA Biol. [Internet]. **18**, 178–193. Available from: https://www.tandfonline.com/doi/full/10.1080/15476286.2019.1695712

[20] Lizarrondo J, Dock-Bregeon A-C, Martino L, Conte MR. (2021) Structural dynamics in the La-module of La-related proteins. RNA Biol. [Internet]. **18**, 194–206. Available from: https://www.tandfonline.com/doi/full/10.1080/15476286.2020.1733799

[21] Billey E, Hafidh S, Cruz-Gallardo I, Litholdo CG, Jean V, Carpentier M-C, et al. (2021) LARP6C orchestrates posttranscriptional reprogramming of gene expression during hydration to promote pollen tube guidance. Plant Cell [Internet]. Available from: https://academic.oup.com/plcell/advance-article/doi/10.1093/plcell/koab131/6276993

[22] Song MH, Aravind L, Müller-Reichert T, O'Connell KF. (2008) The Conserved Protein SZY-20 Opposes the Plk4-Related Kinase ZYG-1 to Limit Centrosome Size. Dev. Cell [Internet]. **15**, 901–912. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1534580708004000

[23] Grimm C, Pelz J-P, Schneider C, Schäffler K, Fischer U. (2020) Crystal Structure of a Variant PAM2 Motif of LARP4B Bound to the MLLE Domain of PABPC1. Biomolecules [Internet]. **10**, 872. Available from: https://www.mdpi.com/2218-273X/10/6/872

[24] Thandapani P, O'Connor TR, Bailey TL, Richard S. (2013) Defining the RGG/RG Motif. Mol. Cell [Internet]. **50**, 613–623. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1097276513004085

[25] Zhou L, Vejlupkova Z, Warman C, Fowler JE. (2021) A Maize Male Gametophyte-Specific Gene Encodes ZmLARP6c1, a Potential RNA-Binding Protein Required for Competitive Pollen Tube Growth. Front. Plant Sci. [Internet]. **12**. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2021.635244/full

[26] Honys D, Twell D. (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol. [Internet]. **5**, R85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15535861

[27] Qin Y, Leydon AR, Manziello A, Pandey R, Mount D, Denic S, et al. (2009) Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. PLoS Genet. [Internet]. **5**, e1000621. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19714218

[28] Gilles LM, Calhau ARM, La Padula V, Jacquier NMA, Lionnet C, Martinant J-P, et al. (2021) Lipid anchoring and electrostatic interactions target NOT-LIKE-DAD to pollen endo-plasma membrane. J. Cell Biol. [Internet]. **220**. Available from:

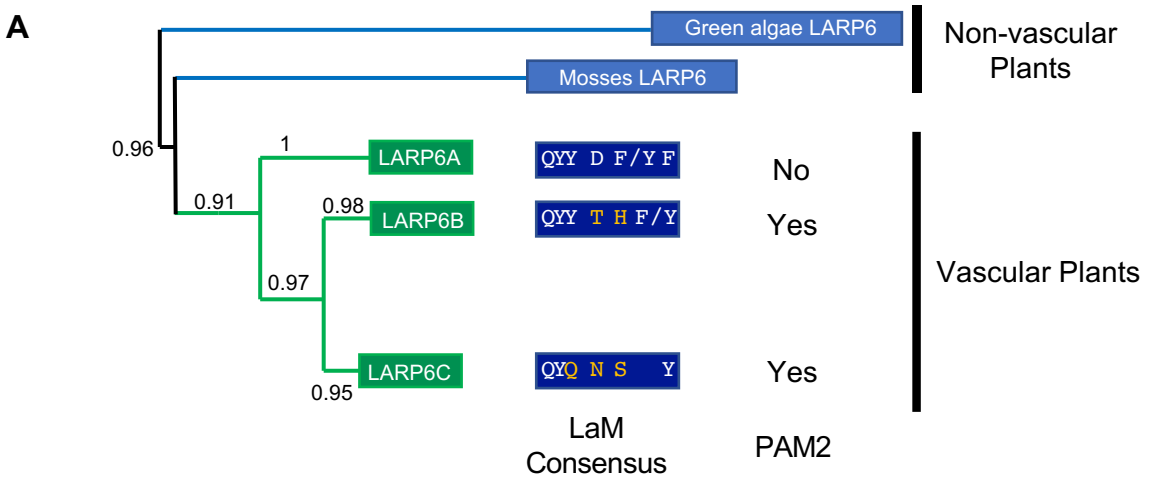
https://rupress.org/jcb/article/220/10/e202010077/212519/Lipid-anchoring-and-electrostatic-interactions

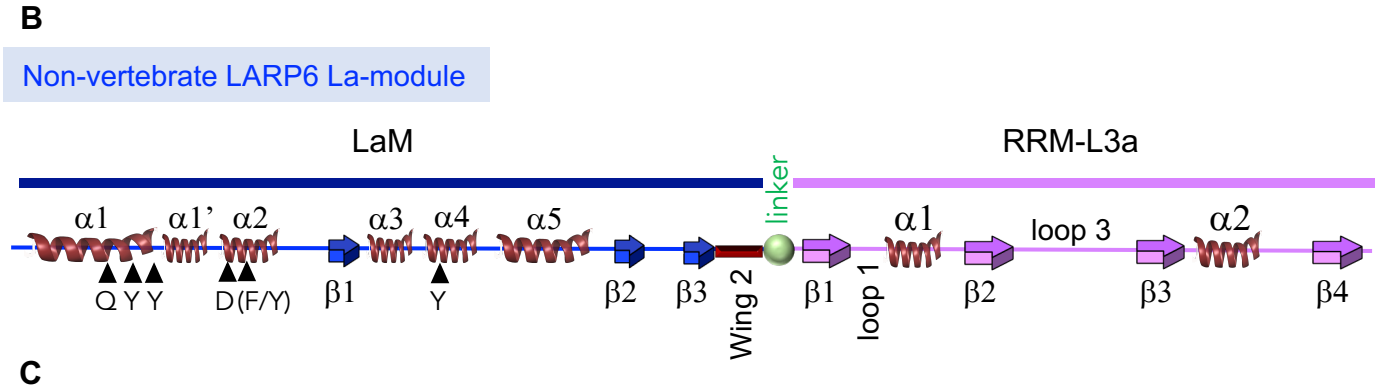
[29] Turner-Bridger B, Caterino C, Cioni J-M. (2020) Molecular mechanisms behind mRNA localization in axons. Open Biol. [Internet]. **10**, 200177. Available from: https://royalsocietypublishing.org/doi/10.1098/rsob.200177

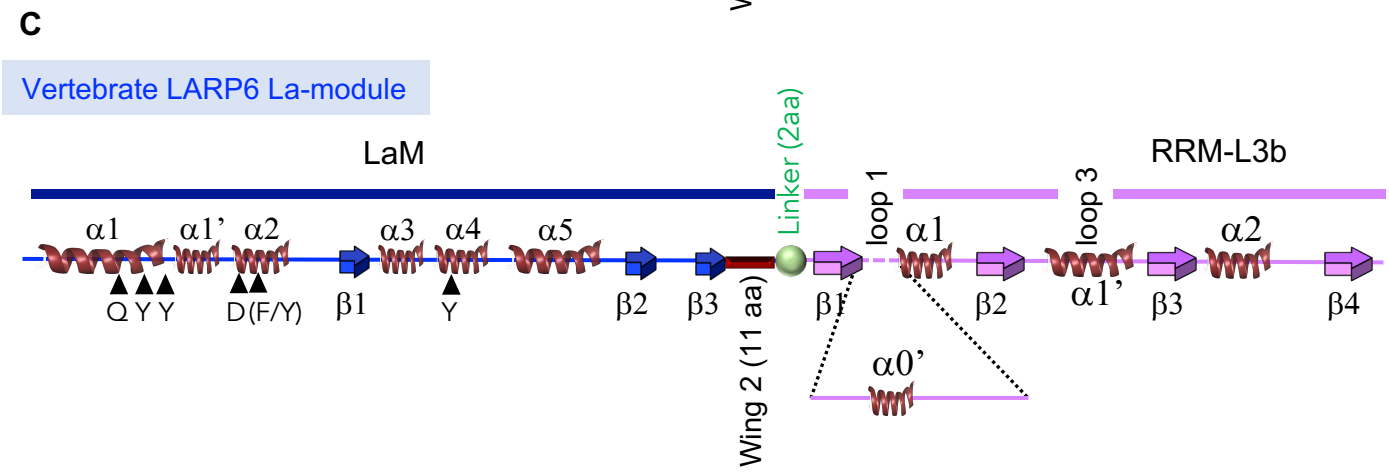
[30] Loraine AE, McCormick S, Estrada A, Patel K, Qin P. (2013) RNA-seq of Arabidopsis pollen uncovers novel transcription and alternative splicing. Plant Physiol. [Internet]. **162**, 1092–109. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23590974

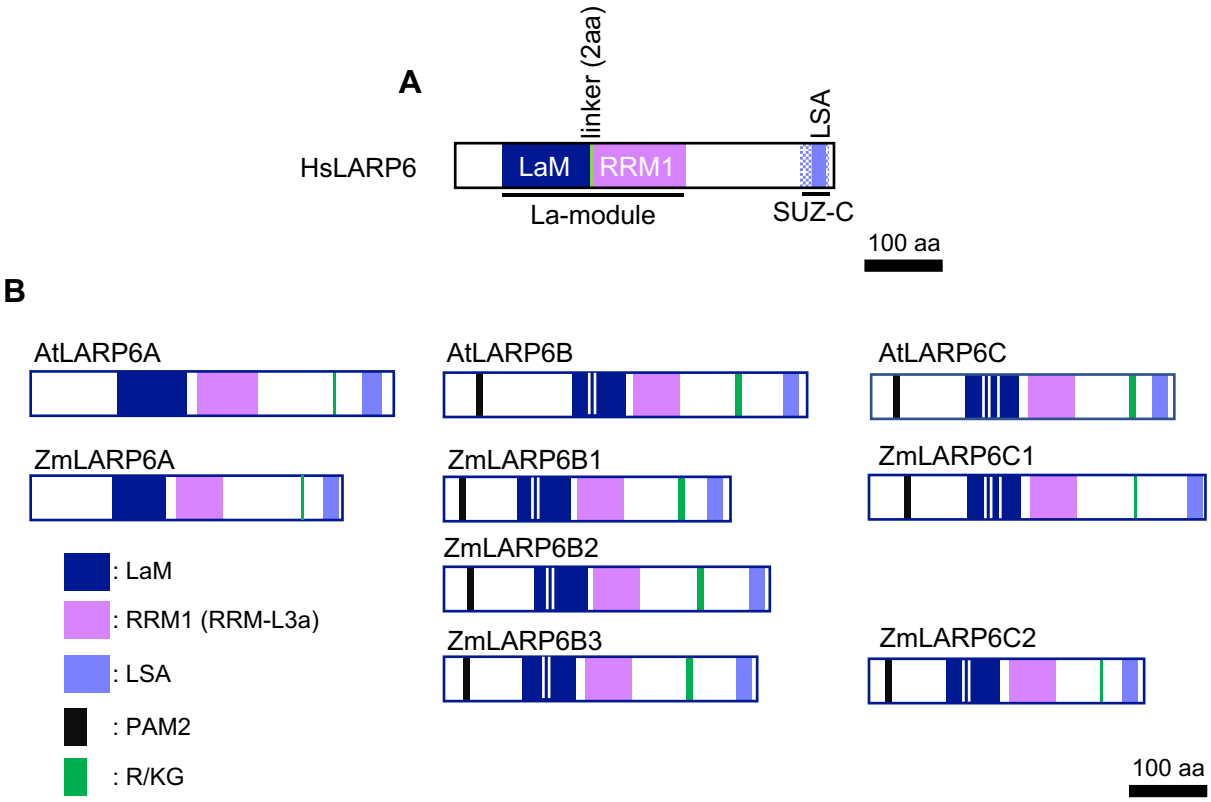
Figure 1: A, Schematic representation of plant LARP6 evolutionary clustering, based on primary sequence comparison of their La-modules. Statistical supports of key nodes are reported from [16]. On the right hand-side of the tree are reported the consensus sequences of the six amino acids known to be involved in RNA binding in genuine La and other LARPs. Orange lettering highlights the non-conservative changes. (**B, C)** Representation of the predicted secondary structure elements carried by the La-modules (LaM and RRM1) from non-vertebrate (**B**, RRM-L3a) and vertebrate (**C**, RRM-L3b) LARP6 proteins. In non-vertebrate LARP6, loop 1 of RRM1 is 7.4 (+/-1.1) aa-long. In non-vascular plant LARP6 and vascular plant LARP6A, loop 3 is 21 nt long while in 6B and 6C loop 3 sizes range from 20 to 25 nt. In RRM1 (RRM-L3b) from vertebrate LARP6, loop 1 is 22 nt long (+/- 4 nt) and was found experimentally to carry an extra helix in HsLARP6 (α0'). Human RRM-L3b also carries an additional α helix, labeled α1' located in loop 3. Finally, wing 2 in HsLARP6 La-module is longer than in HsLa and the linker region shorter (2 versus 11 aa-long) [9].

Figure 2: Schematic representation of the domain organization of Human, *Arabidopsis thaliana* and *Zea maize* LARP6 proteins. A, Schematic representation of non-plant LARP6 protein, using Human protein as model, **B**, Schematic representation of Arabidopsis thaliana and Zea maize LARP6 proteins. Protein representations are organized according to their evolutionary types (6A, 6B or 6C). AtLARP6A (at5g46250), AtLARP6B (at2g43970), AtLARP6C (at3g19090), ZmLARP6A (Zm00001eb152790), ZmLARP6B1 (Zm00001eb037380), ZmLARP6B2 (Zm00001eb411320), ZmLARP6B3 (Zm00001eb112400), ZmLARP6C1 (Zm00001eb298140), ZmLARP6C2 (Zm00001eb043040). Accession numbers for *Zea maize* proteins are from maize genome B73 version 5. Scale bar: 100 amino acids. **C**, Alignments of PAM2, LSA and RG sequences between Maize and Arabidopsis LARP6 proteins.

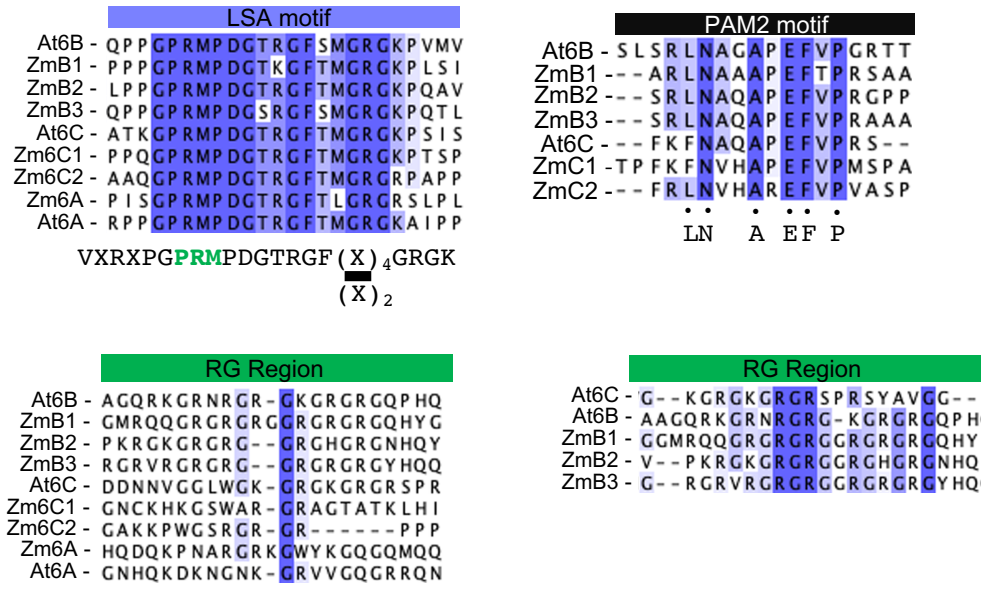

A
Green algae LARP6
Mosses LARP6
Non-vascular Plants
0.96
1
LARP6A
QYY D F/Y F
No
0.91
0.98
LARP6B
QYY T H F/Y
Yes
0.97
Vascular Plants
LARP6C
QYQ N S Y
Yes
0.95
LaM Consensus
PAM2


B
Non-vertebrate LARP6 La-module
LaM
RRM-L3a
linker
α1
α1'
α2
α3
α4
α5
loop 3
α2
Q Y Y
D(F/Y)
β1
Y
β2
β3
Wing 2
β1
loop 1
α1
β2
β3
β4


C
Vertebrate LARP6 La-module
LaM
Linker (2aa)
RRM-L3b
loop 1
loop 3
α1
α2
α1'
α3
α4
α5
Q Y Y
D(F/Y)
β1
Y
β2
β3
Wing 2 (11 aa)
β1
α0'
α1
β2
β3
β4


A
linker (2aa)
LSA
HsLARP6
LaM
RRM1
La-module
SUZ-C
100 aa
B
AtLARP6A
AtLARP6B
AtLARP6C
ZmLARP6A
ZmLARP6B1
ZmLARP6C1
ZmLARP6B2
ZmLARP6B3
ZmLARP6C2
: LaM
: RRM1 (RRM-L3a)
: LSA
: PAM2
: R/KG
100 aa

C


LSA motif
At6B - QPPGPRMPDGTRGFSMGRGKPVMV
ZmB1 - PPPGPRMPDGTKGFTMGRGKPLSI
ZmB2 - LPPGPRMPDGTRGFTMGRGKPQAV
ZmB3 - QPPGPRMPDGSRGFSMGRGKPQTL
At6C - ATKGPRMPDGTRGFTMGRGKPSIS
Zm6C1 - PPQGPRMPDGTRGFTMGRGKPTSP
Zm6C2 - AAQGPRMPDGTRGFTMGRGRPAPP
Zm6A - PISGPRMPDGTRGFTLGRGRSLPL
At6A - RPPGPRMPDGTRGFTMGRGKAIPP
VXRXPGPRMPDGTRGF(X)4GRGK
(X)2
PAM2 motif
At6B - SLSRLNAGAPEFVPGRTT
ZmB1 - ARLNAAAPEFTPRSAA
ZmB2 - SRLNAQAPEFVPRGPP
ZmB3 - SRLNAQAPEFVPRAAA
At6C - FKFNAQAPEFVPRS
ZmC1 - TPFKFNVHAPEFVPMSPA
ZmC2 - FRLNVHARREFVPVASP
LN A EF P
RG Region
At6B - AGQRKGRNRGR-GKGRGRGQPHQ
ZmB1 - GMRQQGRGRGRGGRGRGRGQHYG
ZmB2 - PKRGKGRGRG--GRGHGRGNHQY
ZmB3 - RGRVRGRGRG--GRGRGRGYHQQ
At6C - DDNNVGGLWGK-GRGKGRGRSPR
Zm6C1 - GNCKHKGSWAR-GRAGTATKLHI
Zm6C2 - GAKKPWGSRGR-GR------PPP
Zm6A - HQDQKPNARGRKGWYKGQGQMQQ
At6A - GNHQKDKNGNK-GRVVGQGRRQN
RG Region
At6C - G--KGRGKGRGRSPRSYAVGG--
At6B - AAGQRKGRNRGRG-KGRGRGQPH
ZmB1 - GGMRQQGRGRGRGGRGRGRGQHY
ZmB2 - V--PKRGKGRGRGGRGHGRGNHQ
ZmB3 - G--RGRVRGRGRGGRGRGRGYHQ