[Click here to view linked References](#)

Characterization of the lamin analogue NMCP2 in the monocot *Allium cepa*

Malgorzata Ciska¹, Kiyoshi Masuda², Susana Moreno Díaz de la Espina^{1,*}

¹Cell and Molecular Biology Department, Centre of Biological Research, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain. ² Laboratory of Crop Physiology, Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan.

*Corresponding author: Dr Susana Moreno Díaz de la Espina. smoreno@cib.csic.es; Tel: 34-91 8373112, ext 4257; FAX: 34-91 5360432; ORCID code: 0000-0002-5795-0299

Acknowledgements

We thank M. Carnota, for expert technical assistance. This work was supported by the Spanish Ministry of Science and Innovation [BFU2010-15900] and the Spanish National Research Council (CSIC) [PIE 201020E019].

Abstract

Nuclear lamina organization is similar in metazoan and plants though the latter lack orthologs of lamins, the main components of the metazoan lamina. Current evidence suggests that Nuclear Matrix Constituent Proteins (NMCPs) are the lamin analogues in plants as these proteins share several key features: higher order secondary structure and domain layout, subnuclear distribution, and involvement in the regulation of nuclear shape and size, as well as in higher order chromatin organization. Previously we studied the NMCP family in flowering plants (angiosperms), in which it comprises two phylogenetic groups: NMCP1 and NMCP2. At present, *in silico* information about NMCP proteins in embryophytes is relatively advanced, though very few proteins, most of them of the NMCP1 type, have been extensively studied *in vivo*. We previously characterized the NMCP1 protein in the monocot *Allium cepa*. Here we report the key features of a second protein of this species NMCP2, which presents a conserved sequence and domain layout. Immunofluorescence and immunoelectronmicroscopy evidence co-localization of endogenous AcNMCP2 and AcNMCP1 in the lamina, while western blotting and immuno confocal microscopy reveal a similar pattern of expression and distribution of both NMCP proteins in different root tissues. Our results provide novel insight about endogenous NMCP2 type proteins and complete the characterization of the NMCP family in *A. cepa*, thus advancing the current understanding of these structural proteins constituting the plant lamina.

Key words: *Allium cepa* NMCP2 and NMCP1 proteins, bioinformatic analysis, immunofluorescence and immunoelectronmicroscopy, phylogenetic analysis, plant lamina, protein analysis.

Introduction

The nuclear lamina is a complex protein network that forms a thin layer attached to the inner nuclear membrane and nuclear pore complexes. It provides support for the nuclear envelope (NE) and anchoring sites for chromatin, links the nucleoskeleton (NSK) to the cytoskeleton through the LINC complexes and is involved in the regulation of signalling and gene expression (Simon and Wilson 2011). The lamina is a conserved nuclear structure in eukaryotes (Ciska and Moreno Diaz de la Espina 2014). The metazoan lamina is best characterized and it consists of a polymeric assembly of lamins and numerous lamin-binding proteins that promote its association with the inner nuclear membrane, nuclear pore complexes and chromatin (Goldberg et al. 2008; Gruenbaum and Medalia 2015).

Plants also have a well organized lamina underlying the inner nuclear membrane as demonstrated by field emission scanning electron microscopy (fe-SEM) (Fiserova and Goldberg 2010; Fiserova et al. 2009) and transmission electron microscopy (TEM) after nuclear extraction (Masuda et al.

1993; Minguéz and Moreno Diaz de la Espina 1993; Moreno Diaz de la Espina et al. 1991), which is similar to that of other eukaryotes and vertebrates (Ciska and Moreno Diaz de la Espina 2014). The nuclear lamina is a crowded structure and as many as 660 proteins were identified in the lamina of *Arabidopsis* (*Arabidopsis thaliana*) leaves by LC-MS/MS (Sakamoto and Takagi 2013). Of note, plants lack lamin genes and also those coding for lamin-binding factors, except for SUN domain proteins (Ciska and Moreno Diaz de la Espina 2014; Koreny and Field 2016; Poulet et al. 2017). Nevertheless, the main functions carried out by lamins are effectively performed in plant nuclei. These facts raised the possibility that plant lamina organization is based on plant specific proteins that might be functional homologs of the animal lamins (Ciska and Moreno Diaz de la Espina 2013; 2014). Up to now few proteins of the plant lamina have been characterized, out of which only nuclear matrix constituent proteins (NMCPs) seem to be structural components (Ciska et al. 2013; Masuda et al. 1997; Sakamoto and Takagi 2013). Others are proteins associated to this structure that interact with the NMCP1 proteins such as SUNs (Graumann 2014), KAKU4 (Goto et al. 2014), NEAPs (Pawar et al. 2016), ARP7 as well as the nuclear proteins MYB3 (MYB-type transcription factor 3), C3HC4 RING-finger proteins (SINAT) and the microtubule binding protein (BIM1) recently reported in carrot (Mochizuki et al. 2017).

Due to the lack of sequence similarity with lamins, the structural proteins forming the plant lamina have been sought for taking into consideration their localization within the lamina, a predicted coiled coil structure and certain phenotypic characteristics of mutants (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2013; 2014; Dittmer et al. 2007; Kimura et al. 2010; Masuda et al. 1993; 1997; Sakamoto and Takagi 2013; Wang et al. 2013). Work to date suggests that NMCPs, also known as crowded nuclei (CRWN) in *Arabidopsis* after the phenotype of mutants affecting their function (Dittmer et al. 2007; Wang et al. 2013), are the *bona fide* lamin analogues in plants. NMCPs are highly conserved proteins that present many similarities with lamins: domain layout, subnuclear distribution, insolubility, interaction with SUN proteins and implication in the regulation of nuclear shape and size as well as higher order genome topology and regulation of seed germination (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2013; Dittmer et al. 2007; Graumann 2014; Grob et al. 2014; Kimura et al. 2014; Kimura et al. 2010; Wang et al. 2013; Zhao et al. 2016).

The NMCP protein family has been well characterised in flowering plants (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2014; Kimura et al. 2010; Poulet et al. 2017; Wang et al. 2013). Phylogenetic analysis revealed that the NMCPs have evolved from two progenitor genes: *NMCP1* and *NMCP2*. Dicots contain from two to three NMCP1-type proteins whereas most monocots contain only one. Both monocots and dicots mostly contain a single NMCP2 protein (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2014; Poulet et al. 2017; Wang et al. 2013).

The prediction of secondary structure of NMCP proteins suggested a highly conserved tripartite structure similar to that of lamins with a long central coiled coil domain formed by two segments with highly conserved ends, predicted to dimerize, separated by a linker (Ciska et al. 2013). There are also several conserved regions in the tail domain including a nuclear localisation signal (NLS)

and its adjacent region, an actin binding-like domain, a segment of acidic amino acids, the C-terminus of the protein and several phosphorylation sites (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2013; 2014).

Although the functions of NMCP proteins are not well known, the analysis of mutations in genes coding for the four Arabidopsis NMCP homologs, named CRWN 1-4 (Dittmer et al. 2007; Wang et al. 2013), revealed that quadruple mutants are not viable indicating that these proteins exert essential functions. The analysis of triple mutants showed that at least one CRWN2 or CRWN3 protein is required for viability, while CRWN1 (AtNMCP1) or CRWN4 (AtNMCP2) cannot fulfil the whole range of functions of the family in Arabidopsis (Wang et al. 2013).

The information about endogenous NMCP proteins is scarce and mainly obtained by using antibodies produced against the proteins in the Apiaceae carrot (*Daucus carota*) and celery (*Apium graveolens*) (Kimura et al. 2010; Masuda et al. 1997). In a previous study we characterised the endogenous NMCP1 proteins of several monocot and dicot species. Although the predicted size was similar for both dicot and monocot NMCP1s (145-150 kDa) the proteins from monocots have a smaller size (100 kDa) suggestive of post-translational modifications (Ciska et al. 2013).

Onion (*Allium cepa*), like most monocots, has only two NMCP genes: *AcNMCP1* and *AcNMCP2* that are homologs of Arabidopsis *CRWN1* and *CRWN4*. *AcNMCP1* has been well characterized previously: it is larger than other monocot NMCP1s and resistant to 7M urea and 6M guanidine extraction. *AcNMCP1* is a component of the lamina but also of the internal NSK as demonstrated by immuno confocal microscopy (LSCM) and immuno TEM and its levels are regulated during differentiation in root cells (Ciska et al. 2013).

We present here the characterization of the second protein of the monocot *A. cepa* (*AcNMCP2*) determining its sequence, predicted secondary structure, biochemical properties, subnuclear distribution, co-localisation with *AcNMCP1* and levels in different root tissues. Our results contribute the first exhaustive characterization of an NMCP2 protein, and complete the characterization of NMCPs in this species.

Results and Discussion

Sequence analysis, coiled-coil prediction and phylogeny of *AcNMCP2*

The *AcNMCP2* amino acid sequence, consisting of 880 amino acids, was used in a search for homologs against the proteomes of plant species representing various groups and the matched sequences were aligned using MUSCLE programme. The alignment was used to generate a phylogenetic tree in MEGA7 programme using Maximum likelihood method (Fig1a). In agreement with previous results the whole family is divided into two clusters, one containing NMCP2 and the other NMCP1-type proteins (Ciska et al. 2013; Ciska and Moreno Diaz de la

Espina 2013) (Fig 1a). The sequences found in the moss *Physcomitrella patens*, the spikemoss *Selaginella moellendorffii* and the liverwort *Marchantia polymorpha* which represent the more basal NMCPs, cluster together with NMCP2 proteins (see Fig 1a), suggesting that they share higher sequence similarity with NMCP2 than with NMCP1 and that NMCP2 proteins maintained more ancient characteristics. We included the sequences found in *Amborella trichopoda* a basal angiosperm, which provides important input concerning NMCP evolution in angiosperms. It contains two NMCP genes, one *NMCP1* and one *NMCP2* which cluster with the NMCP1-type and NMCP2 proteins of the rest of angiosperms respectively, suggesting that the separation between the two NMCP types took place at some point in evolution between lycophytes (*S. moellendorffii*) and basal angiosperms (*A. trichopoda*). This is in agreement with the data published by Amborella Genome Project, which confirmed that paralogous gene copies resulted from the epsilon whole genome duplication that took place shortly before the diversification of all living angiosperms (Amborella Genome Project 2013; Jiao et al. 2011).

In all genomes used for the search a single *NMCP2* gene is present and from one to three genes encoding NMCP1-type proteins were found as previously described (Ciska et al. 2013; Ciska and Moreno Diaz de la Espina 2013; Poulet et al. 2017). Brassicaceae, including *A. thaliana* have one gene encoding a NMCP2 protein (CRWN4) and three genes encoding NMCP1-type proteins: CRWN2 and CRWN3, homologous to NMCP3, and CRWN1 homologous to NMCP1. *Allium cepa* as other monocots contain one NMCP protein of each type: AcNMCP1 and AcNMCP2. AcNMCP2 is positioned on the phylogenetic tree between NMCP2s of *Ananas comosus* and *Spirodela polyrrhiza*, two monocot species that are not closely related to grasses (Fig 1a).

Most of the monocot sequences available from genome projects are cereals, which are closely related and have characteristics conserved in Poaceae but not shared with other monocots. For this reason they are not the best models for searching basal traits of NMCP sequences. *Allium cepa* and other non-cereal monocot species, on the other hand, are more distant species and comparison of the sequences from these species is a better source of information on the conserved traits of NMCPs.

AcNMCP2 is predicted to contain three coils positioned at 47- 330 aa; 358-553 aa; 610-730 aa residues. Not only the regions at the ends of predicted coiled-coil domains are highly conserved, as we previously described, but also the length of the coiled-coil domain formed by two to four coiled coil regions is highly conserved among NMCPs (Fig 1b) (Ciska et al. 2013). Coiled coil regions are formed by sequences rich in repetitive heptad pattern of polar and hydrophilic amino acids (Parry 1975). Although protein regions formed by repeated motifs are known to extensively vary in their length across species the length of coiled coil domains is generally highly conserved, even when other domains of the proteins undergo significant changes (Surkont et al. 2015). This conservation does not depend on amino acid sequence variation but represents a conservation of physical size of the domain suggesting that this is crucial for formation of higher order polymers as in the case of lamins and consequently for performing the protein functions (Ben-Harush et al. 2009; Stuurman et al. 1998; Surkont et al. 2015).

NMCP proteins contain a number of highly conserved regions (Ciska et al. 2013). Some of them are present in the basal NMCPs and NMCP1s but were lost in the NMCP2 cluster. This is exemplified by the NLS, followed by the RRYNLRR region or the C-terminus which is preceded by a stretch of acidic amino acids (Fig 1b, S1) that was lost in dicot NMCP2s (Ciska et al. 2013). NMCP2s of *A. trichopoda* and *S. polyrhiza*, as well as AcNMCP2 do not contain the stretch of acidic amino acids nor the conserved region at the C-terminus (Fig 1b). Nevertheless, the region is present in other monocot NMCP2 proteins and is always preceded by a stretch of acidic amino acids (Ciska et al. 2013). Both are present in NMCP of *M. polymorpha*, and the NMCPs in Bryophyta suggesting that it is an ancient trait of NMCP proteins that was lost in dicot NMCP2 and possibly at some point in *Allium cepa*, *S. polyrhiza* and in the lineage leading to *A. trichopoda*. Nevertheless, the tail domain of AcNMCP2 is shorter compared to the other NMCP2 sequences (Fig S1) and the C-terminus of NMCP2s from *S. polyrhiza* and *A. trichopoda* mismatch the corresponding regions in homologous sequences which could suggest that some of the sequence information is missing or is incorrect. The C-terminal acidic region is also present in other proteins. In Cdc34, an ubiquitin-conjugating enzyme, it mediates precise and highly dynamic binding to a basic canyon in the ubiquitin ligase SCF even at low concentration (Kleiger et al. 2009). In yeast, the Pah1 phosphatase, involved in nuclear envelope recycling, also contains an acidic tail that interacts with Nem1-Spo7. This domain is present in all yeast lipin orthologues and its loss disrupts nuclear structure and intermediates functions of Pah1 in membrane biogenesis (Barbosa et al. 2015; Karanasios et al. 2013).

Most of the dicot NMCP2 proteins contain predicted NLS, sometimes two, but *A. trichopoda* and many monocot NMCP2 proteins including AcNMCP2 and *S. polyrhiza* seem to lack a predicted NLS. Kimura et al. (2014) have reported that although NLS motifs are enough for nuclear localization of NMCP1 in *Dacus carota*, another conserved region (RYNLR) that precedes NLS motifs is required for localization in the lamina at the nuclear periphery. This region is conserved across NMCP1-type and is also present in the NMCPs of Bryophyta and *M. polymorpha*. Although no clear NLS nor the conserved RYNLR motif are present in NMCP2 proteins the localization of these to the nuclear periphery was confirmed by fluorescent confocal microscopy using a GFP construct (Sakamoto and Takagi 2013) and by immunofluorescent labelling (Fig 2c', Fig 3c', e') (Kimura et al. 2010) which suggests a different mechanism of nuclear localization operates.

Identification and characterization of endogenous AcNMCP2

The information about endogenous NMCP2 proteins is very limited and restricted to the Apiaceae family (Kimura et al. 2010). To identify the endogenous NMCP2 in onion a polyclonal antibody was raised in chicken against the 161 N-terminal aminoacids of the protein that contains its two most conserved regions (Fig S1). In spite of the high conservation of this region, the antibody proved to be highly specific for AcNMCP2, showing no cross reactivity in western blots with the

other NMCP protein from onion (AcNMCP1) (Fig 2a), which made the antibody a very useful tool for the analysis of NMCP1/NMCP2 co-localization in onion.

Although the molecular weight of AcNMCP2 predicted from its sequence is 103,4 kDa, the endogenous protein migrates in SDS-PAGE gels as a 120 kDa band (Fig 2a), suggesting posttranscriptional modification of the protein or that part of the sequence information is missing (Fig S1). The predicted molecular weights of NMCP2s from dicot and monocot species are similar (110-120 kDa). Also AcNMCP2 has a similar size to carrot (130 kDa) and celery (120 kDa) endogenous NMCP2s (Kimura et al 2010), suggesting a similar size for NMCP2s in dicots and monocots. This contrasts with the case of NMCP1s that despite similar predicted molecular weights (130-140 kDa) show very variable sizes across species. In some cases the sizes of endogenous NMCP1s are higher than predicted (Ciska et al. 2013; Kimura et al. 2010) suggesting incomplete denaturation or posttranslational modifications of the native proteins, but in the case of most monocots they are lower rather suggesting alternative splicing or post-translational modifications (Ciska et al. 2013).

Distribution of AcNMCP2 in the nuclei of meristematic cells

Laser scanning confocal immunofluorescence microscopy (LSCM) with the anti-AcNMCP2 antibody demonstrates a consistent association of AcNMCP2 with the nuclear periphery in nuclei isolated from root proliferative meristematic cell fractions (1-2 mm from the root tip). The protein shows an irregular distribution along the nuclear rim with scarce intranuclear staining (Fig 2c') similar to the distribution of the NMCP2 type proteins in two dicots celery (Kimura et al. 2010) and Arabidopsis (Sakamoto and Takagi 2013). The high specificity of the antibody for AcNMCP2 (Fig 2a) discards a cross detection of AcNMCP1 that localizes in the same subnuclear domain (Fig 2c) and confirms that AcNMCP2 is a component of the nuclear periphery.

AcNMCP2 is a nucleoskeletal component of the lamina and co-localizes with AcNMCP1

To investigate the association of AcNMCP2 with the NSK, the later was isolated by sequential extraction of soluble nuclear protein fractions (Ciska et al. 2013). Immunoblotting with the anti-AcNMCP2 antibody revealed that the protein was mostly present in nuclei and insoluble fractions (F1, F2 and NSK), and that it was resistant to extraction with non-ionic detergent (S1), DNase and low salt concentrations (S2), and only a much reduced fraction of it was extracted with high ionic strength buffer (S3) suggesting a very strong interaction with this structure (Fig 2a) and demonstrating that NMCP2-type proteins are structural components of the NSK. LSCM optical sections of NSK fractions revealed that AcNMCP2 mainly associates with the peripheral NSK, which corresponds to the lamina (Fig 2e') displaying a similar distribution to that found in isolated nuclei (Fig. 2 c') and demonstrating that the protein is an intrinsic component of the lamina.

AcNMCP2 co-localizes with AcNMCP1 in the nuclear periphery (Fig 2c'', d, d') and the insoluble lamina (Fig 2f, f'; Fig S3). Higher magnifications of confocal sections reveal that although AcNMCP1 and AcNMCP2 proteins co-localize in this structure, they seem to form independent localization domains in some places suggesting that the two proteins could form independent networks (Fig 2d, d', f, f'). Preembedding immunoelectron microscopy of isolated NSKs confirms the co-localization of both proteins in the lamina (Fig 2g) while the negative controls showed no labelling (Fig S2).

The high insolubility of AcNMCP1 and AcNMCP2 and their distribution in the extracted lamina suggest that they form the core filamentous network of this structure (Fiserova and Goldberg 2010; Fiserova et al. 2009). The observation of patches containing exclusively AcNMCP1 or AcNMCP2 by LSCM also suggests that the two proteins could form separate networks that would superpose and interact underneath the NE as revealed by stimulated excitation depletion (STED) microscopy for celery NMCP1 and NMCP2 (Masuda et al. unpublished) and for lamins in metazoan (Turgay et al. 2017).

Levels and nuclear distribution of AcNMCP2 in root cells at different stages of proliferation

The level and nuclear distribution of AcNMCP2 was analyzed during the process of proliferation, elongation and early differentiation of root cells by immunoblot and LSCM of nuclear fractions from cells in the meristem (1-2 mm from the root tip), elongation (2-6 mm) and mature (≥ 6 mm) root zones, as well as in the non-proliferating meristem of quiescent roots. Flow cytometry analysis revealed that cells in the elongation and mature zones were mostly non-proliferating, while those in the meristematic zone proliferated and had abundant nuclei with a DNA content ranging between 2C and 4C, therefore corresponding to the S phase. Cells of quiescent meristems were mostly in G1 phase, with no cells in the S-phase (Ciska et al. 2013). In immunoblots, AcNMCP2 was most abundant in meristematic cells, either proliferating or quiescent. Protein levels in the elongation zone remained similar and dramatically decreased in the higher parts of the differentiated zone, with very weak expression in the cells located 18-20 mm from the root tip, which coincides with similar levels of AcNMCP1 (Fig. 3a). Immunofluorescence labelling for confocal microscopy revealed a consistent distribution of AcNMCP2 predominantly at the nuclear periphery in all the root tissues analysed, from non-proliferating meristems of quiescent bulbs (Fig 3c) to proliferating meristems (Fig 2c') and differentiating zone (Fig 3e'). A consistent localization of AcNMCP2 and AcNMCP1 in the lamina of the quiescent, proliferating, and non-proliferating nuclei of the elongation and differentiation zones of the root was also observed (Fig 2d, d', g; Fig 3d, d', f, f'), although in the latter the localization of AcNMCP1 and AcNMCP2 seems to be more intermingled and gaps of the proteins were observed (Fig 3f, f'). These results indicate that the two NMCP proteins of this species (AcNMCP1 and AcNMCP2) are present in the lamina of proliferating cells and also during the whole process of proliferation and early differentiation of roots and that both show a similar pattern of expression and localization in these stages of root development. Gaps, depleted of both AcNMCP1 and AcNMCP2, were frequently observed in the

differentiated nuclei (Fig 3e- f'), as described before for AcNMCP1 (Ciska et al. 2013). The slight differences observed when comparing differentiated nuclei and meristematic ones, as a more intermingled distribution of AcNMCP1 and AcNMCP2 and the presence of gaps depleted of both proteins, suggest that the lamina undergoes structural changes during differentiation in roots.

Conclusions

As other monocots, *A. cepa* has two NMCP genes, one of each type present in Angiosperms: NMCP1 and NMCP2. AcNMCP2 is positioned in the phylogenetic tree between the NMCP2s of non-cereal monocot species as pineapple (*A. comosus*) and duckweed (*S. polyrhiza*). The predicted secondary structure of AcNMCP2 is conserved in relation to monocot NMCP2s and other NMCPs and is predicted to dimerize, but the tail of the protein displays some variations. It lacks the stretch of acidic amino acids and the conserved C-terminus of the protein, both apparently ancient traits of NMCP proteins that may have been lost not only in *A. cepa* but also in the monocot *S. polyrhiza* and in dicot NMCP2s. AcNMCP2 also lacks the predicted NLS in the tail domain, like many other monocots and the basal angiosperm *A. trichopoda*, although its nuclear localization was confirmed by immunofluorescence confocal microscopy suggesting that nuclear localization occurs through a different mechanism. AcNMCP2 is a nucleoskeletal component that co-localizes with AcNMCP1 in the nuclear lamina and shows a similar expression and localization pattern in all root tissues, although both proteins seem to form separate networks. The subtle changes in the distribution of AcNMCP1 and AcNMCP2 revealed by CLSM may suggest that lamina undergoes structural changes during differentiation. These results open the field of the analysis of NMCP proteins in monocots that show a reticulated heterochromatin organization, different from that of the model plant *Arabidopsis* but relatively abundant in angiosperms, and contribute to increase the knowledge of the components of the crowded plant lamina. The characterization of NMCP2 proteins is a prerequisite for a complete understanding of the organization of the lamina in plants.

Materials and Methods

Plant material and culture

Allium cepa L. francesa var. bulbs were grown as described previously (Samaniego et al. 2006). Quiescent meristems were excised from unsoaked bulbs.

Cloning and sequencing of cDNAs for AcNMCP2

Cloning of the cDNA fragments containing a poly(A)-tail of the cDNA was performed using RNA extracted from the callus of *A. cepa* and the PCR primers with the forward sequence of 5'-TCTGCCTTAGCTAAAGCAAGGAA-3' and the reverse sequence of 5'-AAGCAGTGGTATCAACGCAGAGTAC-3'. The reverse primer sequence includes a reverse complement of an adapter DNA added to the 3' end of poly(A) tail during the reverse transcription reaction. Amplification of the DNA fragments containing 5'-end of the cDNA for AcNMCP2 was carried out using a primer with the reverse sequence 5'-TTCGGAGTTCTGCATTAATAAACTCA-3' that includes a reverse complement of the sequence determined for the cDNA amplified by 3'-RACE and the forward primer 5'-TAATATGCCTCTGCCCATCAA(C)17-3' that primes a poly(dG) sequence extended from the 3'-end of the first-stranded cDNA by terminal deoxynucleotidyl transferase in the presence of dG. The amplified DNA was ligated into the pGEM T-Easy vector (Promega) using the TA-cloning method and cloned into *Escherichia coli* DH5 α cells. Plasmids carrying cDNA were cloned and the DNA sequence of the clones was determined. The GenBank/EMBL/DDBJ accession number for AcNMCP2 is LC150620.

Bioinformatics tools

Genome searches were performed using Phytozome v11. (Goodstein et al. 2012), multiple alignments were carried out using Clustal Omega and MUSCLE, NLS searches were done with NucPred and the phylogenetic analysis was performed using MEGA7 (Kumar et al. 2016). The coiled-coil and polymerization state were predicted using MARCOIL (Delorenzi and Speed 2002) and Multicoil2 (<http://groups.csail.mit.edu/cb/multicoil2/cgi-bin/multicoil2.cgi>), respectively. As a threshold 0, 5 probability of coiled coil was used and 14 aminoacids (two repeats of heptad pattern).

Antibody production and synthesis of polypeptides with partial AcNMCP2 sequences

The cDNA fragment encoding the N-terminal 161 amino acids of AcNMCP2 was subcloned into the pEcoli-Cterm 6xHN vector (Clontech) and expressed using an *E. coli* strain DE3 (Agilent). Protein expression was induced with 1 mM isopropyl- β -dthiogalactopyranoside (IPTG) and the bacterial culture kept at 37°C for 4 h. Proteins were extracted from the insoluble fraction of the cell lysate with 8 M urea, 10 mM Na-phosphate buffer (pH 8.0), and 1 mM 2-mercaptoethanol. The AcNMCP2 fragments fused with a 6xhistidine tag were affinity purified on iMAC resin (Biorad). The fraction retained by the resin at 10 mM imidazole was eluted with 300 mM imidazole. The purified protein was dialyzed against 6 M urea in 10 mM Tris-acetate (pH 7.6) and then precipitated by adding 1.5 volumes of acetone. The protein dissolved in PBS containing 0.04% SDS was used for immunization. The anti-AcNMCP2 antibody was commercially generated in chicken by ProPurity Co, Japan.

Isolation of nuclei and nucleoskeleton

Nuclei were isolated as described previously (Samaniego et al. 2006). The isolation of the NSK fraction was achieved by sequential extraction of nuclei with non-ionic detergent, DNase and high salt buffer as previously described (Samaniego et al. 2006) with minor modifications. Freshly isolated nuclei were incubated for 5 minutes with cytoskeleton buffer (CSKB; 10 mM PIPES pH 6.8, 100 mM KCl, 300 mM sucrose, 3 mM MgCl₂, 20 mM DTT, 1 mM EGTA) containing 1% protease inhibitor cocktail (P9599, Sigma-Aldrich) and 0.5% Triton X-100. Next, soluble and membrane associated nuclear proteins were removed by centrifugation at 3000 rpm for 10 min at 4 °C and collected in supernatant (S1). The pellet containing the nuclear insoluble fraction (F1) was next digested with 75 U of Benzonase (P9599 Sigma-Aldrich) in Digestion Buffer (DB; 10 mM PIPES pH 6.8, 50 mM KCl, 50 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 20 mM DTT, 1 mM EGTA, 1% protease inhibitor cocktail, 0.5% Triton X-100) for 1h. Then 1 M (NH₄)₂SO₄ was added slowly to a final concentration of 0.25 M to remove the DNA and DNA-associated proteins (S2), then the sample was incubated for 15 minutes and centrifuged. To the pellet (F2) containing loosely bound proteins, 4 M NaCl was added to a final concentration of 2 M, incubated for 5 minutes and centrifuged. This step releases proteins bound to the NSK (S3) and rendered the insoluble NSK fraction.

PAGE and immunoblotting

Nuclear and NSK pellets were dissolved in lysis buffer (LB: 100 mM Tris-HCl [pH 7.5], 4.5 M urea, 1 M Thiourea, 2% CHAPS, 0.5% Triton X-100, 10 mM DTT) containing protease inhibitor cocktail (Sigma-Aldrich) and 75 U Benzonase (Sigma-Aldrich). The protein content was measured using the modified Bradford Protein Assay (Berkelman 2008), and then protein extracts were mixed with 6x Laemmli buffer and resolved by SDS-PAGE on 8% (w/v) polyacrylamide gels as described previously (Samaniego et al. 2006). The proteins were transferred to nitrocellulose membranes that were probed with anti-AcNMCP2 (1:500), anti AcNMCP1 (1:1000) or anti plant actin 10-B3 (1:1000) (Sigma A0480) antibodies as described previously Samaniego et al. (2006). Positive controls of the anti-NMCP2 primary antibody were performed using increasing concentrations of the immunization polypeptide. MW values were determined using Quantity One 1-D analysis software (BIO-RAD).

Immunofluorescence

Immunofluorescence labelling was performed on suspensions of isolated nuclei or NSK fractions using the anti-AcNMCP2 antibody (1:50) and a secondary Alexa 488 anti-chicken IgY (1:100) as

described (Samaniego et al. 2006). For double labelling experiments the rabbit anti-AcNMCP1 (1:100) and chicken anti-AcNMCP2 (1:50) primary antibodies were used simultaneously for the first incubation and a mix of the corresponding secondary Alexa 488 anti-chicken (1:100) and Alexa 568 anti-rabbit (1:100) antibodies for the second (Ciska and Moreno Diaz de la Espina 2017). Samples were examined with a Leica TCS-SP5 Confocal microscope, using the Leica-confocal software. Co-localization profiles in double labelling experiments were determined using the Leica Application Suite X (v3.0.2) software. Negative controls were performed omitting primary antibodies.

Electron microscopy (EM)

Isolated NSK fractions were fixed in 0.25% formaldehyde (FA) in PBS [pH 7.2] with 0.5% Triton X-100 for 30 min at 4°C, washed in PBS (2 x 10 min) and blocked in blocking buffer (2% BSA in PBS) for 30 min. The samples were subsequently incubated overnight at 4°C with the anti-AcNMCP2 antibody (1:25) and NMCP1 (1:50) in blocking buffer and then washed in PBS containing 0.05% Tween-20 (3 x 15 min). The pellets were incubated for 45 min at room temperature with a 10 nm gold-conjugated secondary anti-chicken IgY antibody (1:100) and 5 nm gold-conjugated anti-rabbit IgG antibody (1:100) in blocking buffer, washed in PBS containing 0.05% Tween-20 (2 x 15 min), fixed in 2% FA in PBS for 1 h at 4°C, washed again in PBS, dehydrated in a graded ethanol series, and embedded in LR White resin (London Resin) as previously described (Ciska and Moreno Diaz de la Espina 2017). Sections were contrasted in aqueous 5% uranyl acetate 30 min and examined in a Jeol 1230 EM at 80 kV. Negative controls were performed omitting primary antibodies during the first incubation.

Conflict of interest: The authors declare that they have not conflict of interest.

Ethical approval: All applicable international, national and institutional guidelines for the care and use of animals were followed.

Figure Legends

Fig. 1 a: Molecular Phylogenetic analysis by Maximum Likelihood method with values of bootstrap consensus. The evolutionary history was inferred using the Maximum Likelihood method based on the JTT matrix-based model. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analysed. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using a JTT model. The analysis involved 80 amino acid sequences. There

were a total of 1752 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. According to the phylogenetic relationship the NMCP proteins can be classified into NMCP1-type and NMCP2 proteins. Monocots express one NMCP1 (empty squares) and one NMCP2 (filled squares) proteins, as well as *Amborella trichopoda* (a basal angiosperm) (NMCP1-empty rhombi, NMCP2- filled rhombi). Dicots express one NMCP2 protein (filled circle) and two or three NMCP1-type proteins (empty circles). The basal NMCP sequences of *Physcomitrella patens*, *Selaginella moellendorffii* and *Marchantia polymorpha* cluster together with NMCP2 proteins suggesting that these share a higher sequence similarity. Sequence accession data are shown in Supplementary Table 1. b: Schematic representation of the predicted coiled-coil regions (MARCOIL) and conserved regions in the tail domain for selected NMCP amino acid sequences of a basal angiosperm and some representative monocots: *A. trichopoda* (Atr1, Atr2), *Ananas comosus* (Anco1, Anco2) *Spirodela polyrhiza* (Spo1, Spo2), *Oryza sativa* (Osa1, Osa2) and *Allium cepa* (AcNMCP1, AcNMP2). Coiled-coil domains consist of two to four coiled coil regions drawn as grey boxes. The total length of the coiled-coil domain is conserved across NMCPs from different species, suggesting the domain size is important for forming higher order structures and fulfilling the protein functions. The sequences were drawn to scale. Conserved regions: nuclear localization signal (arrow pointing up), RYNLR conserved region (arrow pointing down), stretch of acidic amino acids (four vertical stripes) and conserved region at the C-terminus (hexagon)

Fig. 2 Characterization and subnuclear localization of endogenous AcNMCP2 in meristematic nuclei and NSK. (a) Western blot detection of AcNMCP2 in the nuclear (N), insoluble (F1, F2, NSK) and soluble (S1, S2, S3) fractions obtained during extraction of NSKs from isolated nuclei showing the specificity of the antibody. The anti AcNMCP2 antibody recognizes only a 120 kDa band that is present in nuclei and in all the insoluble fractions including the NSK, but not in the corresponding supernatants except for a very faint band observed after high salt extraction. This indicates that only a very small fraction of the protein is soluble in these conditions. (b) Coomassie blue staining of a gel run in parallel displaying the complex protein composition of the fractions. LSCM sections of meristematic nuclei (c') and NSKs (e') after incubation with the anti AcNMCP2 antibody. This shows the distribution of the protein mainly at the nuclear rim corresponding to the lamina (e'). Localization of AcNMCP1 (c) and AcNMCP2 (c') in LSCM sections of isolated nuclei (c to d') and NSKs (e to f') after simultaneous incubation with anti-AcNMCP2 and anti-AcNMCP1 antibodies. Overlays of the green (AcNMCP2) and red (AcNMCP1) channels show co-localization of both proteins at the nuclear rim (d, d') and lamina (f, f'). Overlays of green (AcNMCP2), red (AcNMCP1) and DAPI channels from nuclei show co-localization of both proteins wrapping the condensed chromatin (c''). High magnification of the nuclear periphery showing that the distribution of both proteins in the lamina is not homogeneous. The overlays show micro domains of co-localization (yellow) intermingled with green (AcNMCP2) (green arrow) and red (AcNMCP1) ones (d, d', f, f'). High resolution co-localization of AcNMCP2 and AcNMCP1 in the nuclear lamina (g). Preembedding Immunogoldlabelling of AcNMCP2 (10 nm gold, green arrows) and AcNMCP1 (5 nm gold, red arrows) after simultaneous incubation of NSK

fractions with the corresponding antibodies that shows intermingled labelling of the lamina. Bar= 100 nm

Fig. 3 Expression and distribution of AcNMCP2 and AcNMCP1 in nuclei isolated from different root segments. (a) AcNMCP2 and AcNMCP1 levels detected by Immunoblotting with anti-AcNMCP2 and AcNMCP1 antibodies respectively in nuclear fractions from different root zones: (q) quiescent meristems, (m) proliferating meristems (1-2 mm from tip), elongating zones: (e1: 2-4 mm from tip) and (e2: 4-6 mm from tip) and differentiating zones: (d1:10-12 mm from tip) and (d2: 18-20 mm from tip) (for the localization of the tissues and the corresponding DNA contents and cell cycle stages see Ciska et al. 2013). Anti actin 10-B3 was used as loading control. (b) Coomassie blue stained gel run in parallel that displays the protein composition of the nuclear fractions from different tissues. (c-f') Localization of AcNMCP2 (c', e') and AcNMCP1 (c, e) at the nuclear lamina in LSCM sections of isolated nuclei from quiescent meristems (c to d'), and d1 zones (e-f'). Overlays of the green and red channels demonstrate localization of both proteins at the NE (d, d', f, f') surrounding the condensed chromatin stained with DAPI (c'', e''). The localization pattern of AcNMCP2 (green) and AcNMCP1 (red) in quiescent nuclei is similar to that in meristematic nuclei displaying micro domains of NMCP2 accumulation (green arrows in d'), whereas in differentiated nuclei signals of both proteins seem more intermingled with a higher degree of co-localization and depleted gaps are observed (white arrow) (e to f')

References

- Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants *Science* 342:1241089 doi:10.1126/science.1241089
- Barbosa AD, Sembongi H, Su WM, Abreu S, Reggiori F, Carman GM, Siniosoglou S (2015) Lipid partitioning at the nuclear envelope controls membrane biogenesis *Molecular biology of the cell* 26:3641-3657 doi:10.1091/mbc.E15-03-0173
- Ben-Harush K et al. (2009) The supramolecular organization of the *C. elegans* nuclear lamin filament *Journal of molecular biology* 386:1392-1402 doi:10.1016/j.jmb.2008.12.024
- Berkelman T (2008) Quantitation of protein in samples prepared for 2-D electrophoresis *Methods in Molecular Biology* 424:43-49 doi:10.1007/978-1-60327-064-9_4
- Ciska M, Masuda K, Moreno Diaz de la Espina S (2013) Lamin-like analogues in plants: the characterization of NMCP1 in *Allium cepa* *Journal of experimental botany* 64:1553-1564 doi:10.1093/jxb/ert020
- Ciska M, Moreno Diaz de la Espina S (2013) NMCP/LINC proteins: putative lamin analogs in plants? *Plant signaling & behavior* 8:e26669 doi:10.4161/psb.26669
- Ciska M, Moreno Diaz de la Espina S (2014) The intriguing plant nuclear lamina *Frontiers in plant science* 5:166 doi:10.3389/fpls.2014.00166
- Ciska M, Moreno Diaz de la Espina S (2017) Detection of Endogenous Nuclear Proteins in Plant Cells: Localizing Nuclear Matrix Constituent Proteins (NMCPs), the Plant Analogs of Lamins *Methods Mol Biol* 1560:297-311 doi:10.1007/978-1-4939-6788-9_23
- Delorenzi M, Speed T (2002) An HMM model for coiled-coil domains and a comparison with PSSM-based predictions *Bioinformatics* 18:617-625
- Dittmer TA, Stacey NJ, Sugimoto-Shirasu K, Richards EJ (2007) LITTLE NUCLEI genes affecting nuclear morphology in *Arabidopsis thaliana* *The Plant cell* 19:2793-2803 doi:10.1105/tpc.107.053231
- Fiserova J, Goldberg MW (2010) Relationships at the nuclear envelope: lamins and nuclear pore complexes in animals and plants *Biochemical Society transactions* 38:829-831 doi:10.1042/BST0380829
- Fiserova J, Kiseleva E, Goldberg MW (2009) Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells *The Plant Journal* 59:243-255 doi:10.1111/j.1365-313X.2009.03865.x
- Goldberg MW, Fiserova J, Huttenlauch I, Stick R (2008) A new model for nuclear lamina organization *Biochemical Society transactions* 36:1339-1343 doi:10.1042/BST0361339
- Goodstein DM et al. (2012) Phytozome: a comparative platform for green plant genomics *Nucleic acids research* 40:D1178-1186 doi:10.1093/nar/gkr944
- Goto C, Tamura K, Fukao Y, Shimada T, Hara-Nishimura I (2014) The Novel Nuclear Envelope Protein KAKU4 Modulates Nuclear Morphology in *Arabidopsis* *The Plant cell* 26:2143-2155 doi:10.1105/tpc.113.122168
- Graumann K (2014) Evidence for LINC1-SUN associations at the plant nuclear periphery *PloS one* 9:e93406 doi:10.1371/journal.pone.0093406
- Grob S, Schmid MW, Grossniklaus U (2014) Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila* *Molecular cell* 55:678-693 doi:10.1016/j.molcel.2014.07.009

- Gruenbaum Y, Medalia O (2015) Lamins: the structure and protein complexes Current opinion in cell biology 32:7-12 doi:10.1016/j.ceb.2014.09.009
- Jiao Y et al. (2011) Ancestral polyploidy in seed plants and angiosperms Nature 473:97-100 doi:10.1038/nature09916
- Karanasios E et al. (2013) Regulation of lipid droplet and membrane biogenesis by the acidic tail of the phosphatidate phosphatase Pah1p Molecular biology of the cell 24:2124-2133 doi:10.1091/mbc.E13-01-0021
- Kimura Y, Fujino K, Ogawa K, Masuda K (2014) Localization of *Daucus carota* NMCP1 to the nuclear periphery: the role of the N-terminal region and an NLS-linked sequence motif, RYNLRR, in the tail domain Frontiers in plant science 5:62 doi:10.3389/fpls.2014.00062
- Kimura Y, Kuroda C, Masuda K (2010) Differential nuclear envelope assembly at the end of mitosis in suspension-cultured *Apium graveolens* cells Chromosoma 119:195-204 doi:10.1007/s00412-009-0248-y
- Kleiger G, Saha A, Lewis S, Kuhlman B, Deshaies RJ (2009) Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates Cell 139:957-968 doi:10.1016/j.cell.2009.10.030
- Koreny L, Field MC (2016) Ancient Eukaryotic Origin and Evolutionary Plasticity of Nuclear Lamina Genome biology and evolution 8:2663-2671 doi:10.1093/gbe/evw087
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets Molecular biology and evolution 33:1870-1874 doi:10.1093/molbev/msw054
- Masuda K, Takahashi S, Nomura K, Arimoto M, Inoue M (1993) Residual structure and constituent proteins of the peripheral framework of the cell nucleus in somatic embryos from *Daucus carota* L. Planta 191:532-540 doi:10.1007/BF00195755
- Masuda K, Xu ZJ, Takahashi S, Ito A, Ono M, Nomura K, Inoue M (1997) Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain Experimental cell research 232:173-181 doi:10.1006/excr.1997.3531
- Minguez A, Moreno Diaz de la Espina S (1993) Immunological characterization of lamins in the nuclear matrix of onion cells Journal of cell science 106 (Pt 1):431-439
- Mochizuki R, Tsugama D, Yamazaki M, Fujino K, Masuda K (2017) Identification of candidates for interacting partners of the tail domain of DcNMCP1, a major component of the *Daucus carota* nuclear lamina-like structure Nucleus 8:312-322 doi:10.1080/19491034.2017.1280210
- Moreno Diaz de la Espina S, Barthelme I, Cerezuela MA (1991) Isolation and ultrastructural characterization of the residual nuclear matrix in a plant cell system Chromosoma 100:110-117 doi:10.1007/BF00418244
- Parry DA (1975) Analysis of the primary sequence of alpha-tropomyosin from rabbit skeletal muscle Journal of molecular biology 98:519-535
- Pawar V, Poulet A, Detourne G, Tatout C, Vanrobays E, Evans DE, Graumann K (2016) A novel family of plant nuclear envelope-associated proteins Journal of experimental botany 67:5699-5710 doi:10.1093/jxb/erw332
- Poulet A, Probst AV, Graumann K, Tatout C, Evans D (2017) Exploring the evolution of the proteins of the plant nuclear envelope Nucleus 8:46-59 doi:10.1080/19491034.2016.1236166

- Sakamoto Y, Takagi S (2013) LITTLE NUCLEI 1 and 4 Regulate Nuclear Morphology in *Arabidopsis thaliana* Plant & cell physiology 54:622-633
doi:10.1093/pcp/pct031
- Samaniego R, Jeong SY, de la Torre C, Meier I, Moreno Diaz de la Espina S (2006) CK2 phosphorylation weakens 90 kDa MFP1 association to the nuclear matrix in *Allium cepa* Journal of experimental botany 57:113-124
doi:10.1093/jxb/erj010
- Simon DN, Wilson KL (2011) The nucleoskeleton as a genome-associated dynamic 'network of networks' Nature reviews Molecular cell biology 12:695-708
doi:10.1038/nrm3207
- Stuurman N, Heins S, Aebi U (1998) Nuclear lamins: their structure, assembly, and interactions Journal of structural biology 122:42-66 doi:10.1006/jsbi.1998.3987
- Surkont J, Diekmann Y, Ryder PV, Pereira-Leal JB (2015) Coiled-coil length: Size does matter Proteins 83:2162-2169 doi:10.1002/prot.24932
- Turgay Y et al. (2017) The molecular architecture of lamins in somatic cells Nature 543:261-264 doi:10.1038/nature21382
- Wang H, Dittmer TA, Richards EJ (2013) Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization BMC plant biology 13:200 doi:10.1186/1471-2229-13-200
- Zhao W, Guan C, Feng J, Liang Y, Zhan N, Zuo J, Ren B (2016) The Arabidopsis CROWDED NUCLEI genes regulate seed germination by modulating degradation of ABI5 protein Journal of integrative plant biology 58:669-678
doi:10.1111/jipb.12448

Supplementary data

Supplementary Table 1. Accession data for sequences used for phylogenetic analysis.

Fig. S1 Sequence alignment of selected NMCPs. The sequences of NMCPs of *Allium cepa*: AcNMCP1 (BAM10996) and AcNMCP2 (LC150620) were aligned with NMCP2 proteins from *Ananas comosus* (Anco2), *Oryza sativa* (Osa2) and NMCP from *Marchantia polymorpha* (Mpo) using MUSCLE tool and edited in Jalview. The coiled-coil segments predicted using MARCOIL (Delorenzi and Speed, 2002) are shaded in grey, the predicted NLS in green, the RYNLR conserved region in blue, the stretch of acidic amino acids in red and the conserved region at the C-terminus in orange. The degree of conservation is noted by yellow and brown bars beneath the alignment (generated by Jalview). The region of AcNMCP2 used for antibody production is highlighted by a red dotted box

Fig. S2 a: Positive control of the anti NMCP2 antibody. Western blots of nuclear proteins (N) and increasing amounts of the peptide used for immunization. b to c'': Negative controls of the CLSM immunolabelling analysis of nuclei and NSKs performed omitting incubation of nuclei (b to b'') and NSK (c to c'') with the primary antibodies. Red channel (b, c), green channel (b', c'), DAPI (b'') and DIC (c'') images of the corresponding fields. d: Negative control of the preembedding immunogoldlabelling omitting incubation with the primary antibodies that shows no gold particles

Fig. S3 Co-localization of AcNMCP2 (green channel) and AcNMCP1 (red channel) in NSKs extracted from proliferating meristems, along selected regions of interest (ROIs) and across selected confocal z sections. The line intensity profiles of the green and red channels across the selected lines clearly demonstrate co-localization of both proteins at the lamina

name	species	accession
Smo	Selaginella moellendorffii	XP_002993584.1
AcNMCP1	Allium cepa	BAM10996
AcNMCP2		BAV92747
DcNMCP1	Daucus carota	BAA20407
DcNMCP3		BAN14787
DcNMCP2		BAI67718
Ppa1	Physcomitrella patens	Pp3c2_34830V3.1
Ppa2		Pp3c1_1360V3.1
Mpo	Marchantia polymorpha	Mapoly0160s0011.1
Brst1	Brachypodium stacei	Brast04G088000.1
Brst2		Brast01G134300.1
Anco1	Ananas comosus	Aco012716.1
Anco2		Aco002253.1
CRWN1	Arabidopsis thaliana	AT1G67230.1
CRWN2		AT1G13220.2
CRWN3		AT1G68790.1
CRWN4		AT5G65770.2
Atr1	Amborella trichopoda	evm_27.model.AmTr_v1.0_scaffold00024.349
Atr2		evm_27.model.AmTr_v1.0_scaffold00044.217
Bdi1	Brachypodium distachyon	Bradi3g53047.1
Bdi2		Bradi2g50990.1
Cgr1	Capsella grandiflora	Cagra.1125s0018.1
Cgr3a		Cagra.0613s0001.1
Cgr3b		Cagra.0568s0239.1
Cgr2		Cagra.0917s0008.1
Cpa1	Carica papaya	evm.model.supercontig_179.33
Cpa3		evm.model.supercontig_1.235
Cpa2		evm.model.supercontig_129.57
Cru1	Capsella rubella	Carubv10019693m
Cru3a		Carubv10019731m
Cru3b		Carubv10011605m
Cru2		Carubv10025809m
Csa1	Cucumis sativus	Cucsa.238180.1
Csa3		Cucsa.280830.1
Csa2		Cucsa.103490.1
Csi1	Citrus sinensis	orange1.1g048767m
Csi3		orange1.1g000847m
Csi2		orange1.1g002268m
Esa1	Eutrema salsugineum	Thhalv10018034m
Esa3a		Thhalv10019562m
Esa3b		Thhalv10006601m
Esa2		Thhalv10003578m
Fve1	Fragaria vesca	mrna10337.1-v1.0-hybrid
Fve3		mrna03889.1-v1.0-hybrid
Fve2		mrna05584.1-v1.0-hybrid
Mgu1	Mimulus guttatus	Migut.H00533.1
Mgu3		Migut.L01107.1
Mgu2		Migut.H00316.1
Mtr1	Medicago truncatula	Medtr7g018610.1

Mtr3		Medtr6g015285.1
Mtr2		Medtr4g097580.1
Osa1	Oryza sativa	LOC_Os02g48010.1
Osa2		LOC_Os01g56140.1
Pha1	Panicum hallii	Pahal.A03119.1
Pha2		Pahal.E01502.1
Ppe1	Prunus persica	Prupe.3G240800.1
Ppe3		Prupe.1G303700.1
Ppe2		Prupe.6G214100.1
Ptr1	Populus trichocarpa	Potri.017G111400.1
Ptr3		Potri.008G114800.1
Ptr2		Potri.012G034300.1
Pvu1	Phaseolus vulgaris	Phvul.008G020600.1
Pvu3		Phvul.003G025500.1
Pvu2		Phvul.003G232200.1
Rco1	Ricinus communis	29673.m000916
Rco3		29738.m001028
Rco2		29825.m000318
Sbi1	Sorghum bicolor	Sobic.004G264300.1
Sbi2		Sobic.003G308200.1
Sit1	Setaria italica	Seita.1G292700.1
Sit2		Seita.5G331200.1
Spo1	Spirodela polyrhiza	Spipo18G0020000
Spo2		Spipo3G0109200
Svi1	Setaria viridis	Sevir.1G297800.1
Svi2		Sevir.5G335100.1
Tca1	Theobroma cacao	Thecc1EG019537t1
Tca3		Thecc1EG011885t1
Tca2		Thecc1EG000864t1
Zma1	Zea mays	GRMZM2G015875_T01
Zma2		GRMZM2G320013_T01

Figure S1