Title:

Identification and characterization of conserved and divergent genes encoding the nuclear envelope LINC complex in maize (*Zea mays* L.).

Running Title:

Maize LINC complex

Authors and Affiliations:

Gumber, Hardeep K.[1]; McKenna, Joseph F.[2]; Estrada, Amado L.[1]; Tolmie, Andrea F.[2] Graumann, Katja[2]; Bass, Hank W.[1]

1. Department of Biological Science, Florida State University, Tallahassee, FL, USA, 32306-4295
2. Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK, OX30BP.

Corresponding Author:

bass@bio.fsu.edu

Key words:

LINC, SUN, KASH, nuclear envelope, maize

SUMMARY STATEMENT

Genes encoding maize candidates for the core LINC and associated complex proteins have been comprehensively identified with functional validation by one or more assays for several of the KASH genes.

ABSTRACT

The LINC (Linker of Nucleoskeleton to Cytoskeleton) complex serves as an essential multi-protein structure spanning the nuclear envelope. It connects the cytoplasm to the nucleoplasm and functions to maintain nuclear shape and architecture, as well as regulates chromosome dynamics during mitosis and meiosis. Knowledge of LINC complex composition and function in the plant kingdom is relatively limited, especially in the monocots which include the cereal grains and other grass species. We identified and classified 22 genes encoding candidate LINC complex and associated proteins in maize through bioinformatic and biochemical approaches. Representative KASH candidates were functionally validated in one or more assays including nuclear envelope localization, native or heterologous co-immunoprecipitation with antisera for ZmSUN2, and fluorescence recovery after photobleaching. These findings support a summary working model of the entire maize LINC and associated proteins complex with components found to be conserved across eukaryotes, unique to plants, or highly divergent and grass-specific. This model contributes a new experimental system for the cell biology of the nuclear envelope and new opportunities for future studies of the LINC complex in a model crop species.

47 List of Symbols and Abbreviations:

48	CDS	coding sequence
49	CRWN1	CROWDED NUCLEI 1
50	FRAP	Fluorescence Recovery After Photobleaching
51	INM	inner nuclear membrane
52	KASH	Klarsicht/ANC-1/Syne-1 homology
53	LINC	linker of nucleoskeleton and cytoskeleton
54	MLKP	Maize LINC KASH AtWIP-like
55	MLKG	Maize LINC KASH Grass-specific
56	MLKS	Maize LINC KASH AtSINE-like
57	MLKT	Maize LINC KASH AtWIT-like
58	MKAKU4	Maize KAKU4-like
59	NE	nuclear envelope
60	NET	nuclear envelope transmembrane protein
61	NCH	NMCP/CRWN-Homologous
62	NEAP	nuclear envelope-associated protein
63	ONM	outer nuclear membrane
64	SUN	Sad1/UNC-84
65	WIP	WPP domain-interacting protein
66	WIT	WPP domain-interacting tail-anchored protein

67 **INTRODUCTION**

68 In eukaryotic cells, the nuclear envelope (NE) is a structural hallmark that
69 encapsulates the biparentally-inherited genetic material, the chromosomes. The NE is
70 comprised of inner and outer nuclear membranes, with embedded nuclear pore
71 complexes (NPC) and a variety of nuclear envelope transmembrane proteins (NETs).
72 Two NETs, Sad1/UNC-84 (SUN), and Klarsicht/ANC-1/Syne-1 homology (KASH)
73 domain proteins, interact in the perinuclear space to form the core of an evolutionarily
74 conserved Linker of Nucleoskeleton and Cytoskeleton (LINC) complex which spans the
75 NE (Crisp et al., 2006). LINC complexes carry out many functions ranging from
76 organizing the shape and position of the nucleus as a mobile and pliable organelle to
77 direct mechanotransduction to effects on chromatin structure and dynamics in somatic
78 and meiotic cells (Kim et al., 2015; Kracklauer et al., 2013; Meier, 2016; Meier et al.,
79 2017; Razafsky and Hodzic, 2015; Starr and Fridolfsson, 2010; Tamura et al., 2015;
80 Tapley and Starr, 2013). The biological importance of NE functions is reflected in part
81 through the wide range of NE-defective developmental disorders such as laminopathies
82 or envelopathies (Burke and Stewart, 2014; Fridkin et al., 2009; Janin et al., 2017;
83 Razafsky and Hodzic, 2015).

84 The core components of the LINC complex are the inner nuclear membrane
85 (INM) SUN-domain proteins (Hagan and Yanagida, 1995; Malone et al., 1999) and the
86 outer nuclear membrane (ONM) KASH-domain proteins (Starr and Han, 2002). The
87 SUN-domain proteins are known to interact with chromatin, nuclear lamins, or lamin-like
88 proteins (Haque et al., 2006; Hodzic et al., 2004; Zhou et al., 2015b). The KASH-
89 domain proteins are known to interact directly or indirectly with cytoskeletal structures

90 including microfilaments, microtubules, intermediate filaments, microtubule organizing
91 centers, and organelles (Luxton and Starr, 2014; Starr and Fridolfsson, 2010). LINC
92 complexes are highly conserved in all eukaryotes yet many of the components have
93 evolved rapidly or recently, limiting the power of sequence conservation to identify
94 functional homologs.

95 The body of knowledge of the LINC complex comes primarily from studies in
96 opisthokonts, especially metazoans and yeast, with some but limited information for
97 other eukaryotes, including plants. Knowledge of plant LINC complex components,
98 especially in Arabidopsis, has recently began to expand (Meier, 2016; Meier et al.,
99 2017; Poulet et al., 2017a; Tiang et al., 2012; Zhou et al., 2015a). The first SUN
100 domain protein recognized in plants, OsSad1, was found as part of nuclear proteomic
101 study in rice (Moriguchi et al., 2005). Its relationship to the *Saccharomyces pombe* Sad1
102 was confirmed by localization of OsSad1-GFP fusion protein to the nuclear periphery in
103 onion epidermal cells (Moriguchi et al., 2005). Subsequently, two classes of SUN
104 domain proteins were found to be present in plants; first, the CCSD/C-terminal SUN
105 group, in which the SUN domain is near the C-terminus and second, the PM3/Mid-SUN
106 group, in which the SUN domain is centrally located (Graumann and Evans, 2010;
107 Graumann et al., 2014; Murphy et al., 2010; Oda and Fukuda, 2011). More recently, a
108 large number of SUN domain proteins have been identified in many plant species by
109 homology searches (reviewed by Meier, 2016; Meier et al., 2017; Poulet et al., 2017a).

110 The C-terminal SUN proteins are similar to those in animals, but the mid-SUN
111 proteins are less well studied as a group despite their known occurrence in plants
112 (Poulet et al., 2017a), yeast (Friederichs et al., 2011), Dictyostelium (Shimada et al.,

113 2010) and mice (Sohaskey et al., 2010). Among the conserved roles that SUN proteins
114 play in plants are those involving the maintenance of the shape and size of the nucleus
115 (Graumann et al., 2010; Graumann et al., 2014), plant growth (Graumann et al., 2014),
116 and meiotic chromosome behavior (Murphy et al., 2014; Varas et al., 2015).
117 The first plant KASH proteins to be described were the WPP-domain-interacting
118 proteins (WIPs). The WIPs are plant-specific, outer nuclear envelope associated
119 proteins originally discovered as RanGAP-NE anchoring proteins (Xu et al., 2007) and
120 later shown to bind to SUN and possess KASH-like features (Meier et al., 2010; Zhou et
121 al., 2012). A related group of ONM LINC-associated plant proteins are the WPP-
122 interacting tail-anchored (WIT) proteins (Zhao et al,. 2008), which bind directly to WIPs,
123 indirectly to SUN, and also have WPP-interacting domains. WITs possess the LINC-like
124 activity of connecting the NE to the actin cytoskeleton through their association with
125 plant specific Myosin XI-i (Tamura et al., 2013). In addition to the WIPs and WITs, two
126 other categories of ONM KASH proteins have been described in plants, the SUN-
127 Interacting Nuclear Envelope proteins (SINEs) and the TIR-KASH (TIKs), both of which
128 are NETs (Graumann et al., 2014; Zhou et al., 2014).
129 There are currently recognized three plant-specific categories of genes encoding
130 proteins in the INM or at the intranuclear periphery that interact with SUN or each other
131 to associate with the LINC complex. These encode the INM Nuclear Envelope
132 Associated Proteins (NEAPs) and nucleoplasmic proteins, including Nuclear Matrix
133 Constituent Proteins (NMCPs, also called CRWNs) and the AtCRWN-binding
134 nucleoplasmic protein AtKAKU4s (Ciska et al., 2013; Dittmer and Richards, 2008; Goto
135 et al., 2014; Masuda et al., 1997; Pawar et al., 2016; Wang et al., 2013). While NEAPs

136 and CRWNs can be directly associated with LINC complexes by binding to C-terminal
137 SUN proteins, KAKU4 is indirectly associated with LINC complexes via interaction with
138 CRWNs (Goto et al., 2014). These interconnected proteins are thought have roles in
139 chromatin architecture and nuclear morphology, placing the LINC complex squarely in
140 genetic and genomic response pathways in plant cells (Poulet et al., 2017b).
141 In this study, we have identified, named and tested some of the core LINC
142 complex components in maize, an important agricultural and economic species with
143 historic contributions to genetics, cytogenetics, epigenetics, and meiosis (reviewed by
144 Nannas and Dawe, 2015). We include comparisons across phyla ranging from single
145 cellular chlorophyte to advanced land plants to highlight both conserved and divergent
146 maize LINC complex components.
147

148 ## RESULTS

149 **Identification of maize genes encoding LINC or LINC-associated components**

150 We developed a working model, shown in Figure 1, for the maize LINC complex
151 and associated proteins as identified through bioinformatic or experimental evidence.
152 The proteins are depicted as residing in one of the four cellular locations, the cytoplasm,
153 the ONM of the NE, the INM of the NE, or the nucleoplasm. Despite the considerable
154 difficulty in assigning certain orthology for all of the maize LINC components, we have
155 established the following nomenclature for these candidate maize LINC genes. The
156 genes encoding KASH-like and associated ONM proteins are named with the common
157 prefix of *MLK* for *Maize LINC KASH* followed by a one letter designation to delineate the
158 four KASH-related subgroups in maize: *G* for Grass-specific, *S* for *AtSINE*-like, *P* for

159 *AtWIP*-like, or *T* for *AtWIT*-like. The genes encoding the INM LINC proteins fall into two
160 groups, the previously published core LINC SUN domain proteins (Murphy et al., 2010),
161 and the genes designated *MNEAP* for Maize *AtNEAP*-like. Finally, the genes encoding
162 the nucleoplasmic LINC-associated proteins are designated as *NCH* for NMCP/CRWN
163 homologs or *MKAKU4* for *Maize KAKU4-like*. In summary the 22 known or candidate
164 core LINC and associated proteins in maize are comprised of 10 MLKs, 5 SUNs, 3
165 NEAPs, 2 NCHs, and 2 MKAKU4s. For all of these, the gene IDs in maizeGDB, PLAZA
166 gene, and UniPROT are summarized in Table 1 and diagrammed with their protein
167 domain structures in Figure 2. Of these, the KASH and SUN proteins are the most
168 similar to their animal counterparts, whereas the MLKGs, MLKTs, MNEAPs, MKAKU4s
169 appear to be more characteristic of plants.
170

171 **The MLK genes: ten candidates for maize ONM core LINC genes**

172 A hallmark of the KASH domain proteins is that they have limited sequence
173 conservation and are not reliably recognized by sequence similarity searches. Instead,
174 they contain a conserved structure consisting of a C-terminal transmembrane domain,
175 and a short periplasmic tail with characteristic terminal residues including a penultimate
176 proline followed by a threonine or serine. Only a few plant genes with similarity to
177 animal KASH genes have been identified so far (Zhou et al., 2012; Zhou et al., 2014).
178 To find all of the potential maize LINC ONM components, we first searched the maize
179 proteome (B73 AGPv4), from ensembl.gramene.org (Jiao et al., 2017) for KASH
180 proteins with the following characteristic features; (1) a TM domain at the C-terminus
181 followed by 9-40 amino acids that could extend into the perinuclear space, and (2)

182 terminal four residues of [DTVAMPLIFY], [VAPIL], P, and T (Poulet et al., 2017a; Zhou
183 et al., 2014). This search led to the identification of a total of eight genes encoding
184 maize KASH-like candidates, four with homology to AtWIPs (*MLKP1* – *MLKP4*), two
185 with homology to AtSINEs (*MLKS1* and *MLKS2*) and two found only in members of the
186 grass family, Poaceae (*MLKG1* and *MLKG2*). We next searched the maize proteome
187 for WIT homologs using AtWIT1 and AtWIT2 protein sequences as queries (Zhao et al.,
188 2008) and found two maize genes, *MLKT1* and *MLKT2*, both encoding proteins with the
189 AtWIT-like TM domain at the C-terminus and multiple coiled-coil (CC) domains. Given
190 the rapid divergence of LINC gene sequences and the lack of distinct and diagnostic
191 domains, and gene content differences between maize cultivars (Hirsch et al., 2016), it
192 is possible that maize has even more MLKT genes not readily detected using these
193 criteria.

194 The four MLKP (AtWIP-like) genes belong to the PLAZA gene family designated
195 HOMO4M002181, and encode proteins with multiple coiled-coil domains (MLKP1-
196 MLKP4, Fig 2). The previously identified AtWIP proteins interact with AtWITs and
197 RanGAP1 and exhibit coiled-coil domain-dependent homo- and hetero-multimerization
198 (Zhao et al., 2008; Zhou et al., 2012). The MLKP proteins are therefore predicted to
199 interact with themselves, each other, ZmSUN proteins, and RanGAP proteins with WPP
200 domains. The two MLKS (AtSINE-like) genes belong to PLAZA gene family designated
201 HOMO4M002474, and encode proteins with armadillo (ARM) repeats (MLKS1- MLKS2,
202 Fig. 2). ARM domains have been reported in a variety of proteins across plant and
203 animal kingdoms as actin-binding and protein-protein interaction domain (Coates,
204 2003). The MLKS proteins are predicted, therefore, to interact and colocalize with actin,

205 as has been demonstrated for AtSINE1 (Zhou et al., 2014). The two MLKG (grass-
206 specific) genes lack both CC and armadillo domains and have been assigned to
207 different PLAZA gene families - HOMO4M009187 for *MLKG1* and HOMO4M008346 for
208 *MLKG2*.

209 Beyond the eight candidate MLK genes that we predict to encode maize ONM
210 core LINC complex proteins, we also found four other maize genes that met some of
211 our criteria. For instance, some encoded proteins with the terminal PT residues, but did
212 not satisfy the single C-terminal TM criterion. It remains unknown whether these genes
213 (B73v4 gene IDs Zm00001d024832, Zm00001d036395, Zm00001d005590 and
214 Zm00001d017144) encode actual variant LINC proteins or simply resemble LINC
215 proteins.

216

217 **The eight maize INM LINC genes; five SUN and three NEAP genes**

218 The five genes encoding maize SUN-domain proteins belong to two subfamilies
219 distinguished by the location of the SUN domain (Murphy et al., 2010). They are not
220 newly identified in this study, but are included in Table 1 for the sake of completion,
221 along with their most recent B73v4 gene IDs. ZmSUN1 and ZmSUN2 proteins have the
222 canonical C-terminal SUN domain, whereas the ZmSUN3, ZmSUN4,and ZmSUN5 have
223 a central SUN domain that is characteristic of, but not limited to plants (Graumann et al.,
224 2014). Despite the conservation of SUN-domain proteins as a group, and their
225 assignment to one of two subfamilies, direct assignment of orthology for individual SUN
226 genes across taxonomic groups such as eudicots (Arabidopsis) versus monocots
227 (maize) has yet to be established.

228 Another major family of INM LINC proteins are the higher plant-specific proteins
229 called Nuclear Envelope-Associated Proteins, the NEAPs (reviewed by Groves et al.,
230 2018; Meier et al., 2017; Poulet et al., 2017a). NEAPs exist as a multigene family, first
231 discovered in Arabidopsis and shown to interact with SUN-domain proteins (Pawar et
232 al., 2016). Genetic analysis of AtNEAPs showed that they affect several phenotypes,
233 including primary root growth, nuclear morphology, and chromatin organization (Pawar
234 et al., 2016) . NEAPs have three protein structural features, the presence of extensive
235 coiled coil domains, nuclear localization sequences, and C-terminal TM domains. We
236 used these criteria to identify three maize NEAP genes (Table 1, Fig. 2), *MNEAP1*,
237 *MNEAP2*, and *MNEAP3*. Interestingly, we observed partial homology to N-terminus of
238 structural maintenance of chromosomes domain (SMC_N, PFAMID PF02463) proteins
239 in the MNEAPs, further implicating them in DNA and chromosomal functions.
240 The MNEAP1 and MNEAP2 proteins both end with a characteristic C-terminal
241 hydrophobic patch and two basic terminal residues KR (Pawar et al., 2016). These two
242 proteins share 92% sequence identity and are very similar to the protein encoded by
243 one of the two sorghum NEAPs genes (sobic.009G033800). The *MNEAP1* and
244 *MNEAP2* most likely arose from the genome duplication event that occurred following
245 the divergence of maize from sorghum, ~ 11 MYA (Gaut et al., 2000). *MNEAP3* fits our
246 initial criteria but is otherwise quite different in many respects including a unusually
247 large number of splice variant transcript models and the lack of KR terminal residues.
248 Consequently, *MNEAP3* may be a complex locus or alternatively, pseudogene no
249 longer under evolutionary pressure to maintain its structure.
250

251 **The four maize nucleoplasmic LINC-associated genes; two *NCH* and two *KAKU4***
252 **genes**

253 In mammals, the SUN-domain proteins of the LINC complex interact with the
254 nuclear lamina which is located at the periphery of the nucleoplasm and made up of
255 lamin proteins. Plant nuclei lack laminA/B orthologs, but they do have a lamin-like
256 meshwork that is thought to consist of lamin-like proteins known as NMCP/CRWN as
257 well as CRWN-binding nuclear proteins, KAKU4, (Ciska et al., 2013; Goto et al., 2014).
258 These proteins are plant-specific, collectively thought to comprise plant nuclear
259 architecture proteins, and are associated with the LINC complex through CRWN-SUN
260 interactions (Goto et al., 2014).

261 Using AtCRWN and DcNMCP protein sequences as queries, we identified two
262 maize genes, *NCH1* and *NCH2*, representing the two major clades of NMCP/CRWN
263 genes (Ciska et al., 2013; Wang et al., 2013). The maize *NHC1* is most similar to
264 *DcNMCP1*, *DcNMCP3*, and *AtCRWN1*. The genes in this group are members of
265 PLAZA gene family HOM04M004640 (Table 1). The maize *NCH2* gene is most similar
266 to *DcNMCP2*, *AtCRWN2*, *AtCRWN3*, and *AtCRWN4*. The genes in this group are
267 members of the PLAZA gene family HOM04M003564. Both NCH1 and NCH2 proteins
268 share several features with their plant homologs. These features include large proteins
269 (~100 kDa or greater), multiple coiled-coil domains over the majority of the protein,
270 multiple CDK1 phosphorylation sites, and a nuclear localization sequence near the C-
271 terminus (Fig. 2).

272 Using AtKAKU4 protein sequence similarity searches, we identified two highly
273 similar maize KAKU4-like genes, *MKAKU41* and *MKAKU42*. These two genes and

274 AtKAKU4 are all members of the PLAZA gene family HOM04M004991. KAKU4 and
275 KAKU4-like genes appear to be limited to flowering plants and encode proteins that
276 have been described as candidate lamina components (Goto et al., 2014). As a group,
277 they may be remotely associated with the LINC complexes but are included here
278 because of the known interaction of AtKAKU4 with AtCRWN proteins as well as the
279 genetic evidence that implicates them in LINC functions and nuclear architecture (Goto
280 et al., 2014). Aside from a NLS near the N-terminus, the MKAKU4 proteins do not
281 contain any notable structural protein domains or motifs (Fig. 2).
282

283 **A mix of ancient and derived genes encode components of the maize LINC**
284 **complex**

285 Given that the time of appearance and the evolutionary conservation of genes
286 often reflects functional properties, we examined the maize LINC candidate genes and
287 their counterparts in a variety of plant species as summarized in Figure 3. Recent
288 phylogenetic analyses of plant LINC genes or sub-families therein have been published
289 (Ciska et al., 2013; Poulet et al., 2017a; Zhou et al., 2014), with emphasis on
290 Arabidopsis and other eudicot species, but did not include the latest PLAZA gene
291 families or maize B73v4 annotations (Jiao et al., 2017; Van Bel et al., 2018). The
292 species listed (Fig. 3) are arranged in increasing evolutionary distance from maize
293 starting with sorghum and extending to the unicellular green alga, *Chlamydomonas*
294 *reinhardtii*.

295 In plants, the most highly conserved members of the core LINC and associated
296 complex belong to gene families that encode the C-terminal and mid-SUNs, NCH2 and

297 the AtSINE-like KASH proteins MLKS1 and MLKS2). These gene families are present
298 in all the plant species listed (shaded blue or green, Fig. 3). Curiously, of all the genes
299 listed, only the mid-SUN genes are listed as present in Chlamydomonas PLAZA gene
300 families. The next most highly conserved plant LINC gene families encode the MNEAP
301 and NCH1 proteins. They are present in seed-bearing plants (shaded tan, Fig. 3),
302 which includes pine, a conifer species. The third most highly conserved plant LINC
303 gene families encode MLKP, MLKT and MKAKU4 proteins, all of which are found in
304 flowering plants, the angiosperms (shaded pink, Fig. 3). In contrast, the most recent to
305 appear in evolutionary time are the gene families that encode the MLKG proteins,
306 present in the grass family within the monocots (shaded yellow, Fig. 3). The biological
307 significance of the timing of evolutionary appearance remains to be determined, but we
308 generally expect that the more widely distributed gene families have more ancient or
309 basal functions, whereas the more recent gene families could have resulted from
310 duplication followed by either regulatory divergence, such as tissue-specific expression,
311 or gain of new biological functions.
312

313 **Maize LINC genes exhibit tissue-specific and developmental co-expression**
314 **patterns**

315 We analyzed the expression patterns of the maize LINC candidate genes as
316 shown in Figure 4 in order to examine them for evidence of expression and potential
317 functional groupings that should be reflected in their differential or co-expression
318 patterns. For this, we use the expanded maize gene expression atlas which contains
319 transcriptome data from numerous tissues broadly sampled across most of the

320 developmental stages of the maize life cycle (Stelpflug et al., 2016). For comparison,
321 we included three housekeeping genes (HKG, Fig. 4A), *CDK1*, *RPN*, and *SGT1* (Lin et
322 al., 2014). We excluded *MNEAP3*, which lacked a gene model when the atlas was
323 produced, and *MLKG1*, which had zero reads in all fifteen tissues selected for this
324 study. Despite their exclusion, these two are presumed to be real genes and not
325 pseudogenes because of the expression of their corresponding transcripts detected in
326 one or more tissues (Chettoor et al., 2014; Walley et al., 2016; Wang et al., 2016; Xu et
327 al., 2012).

328 Most of the candidate maize core LINC and LINC-associated genes showed a
329 pattern of broad expression across multiple tissues, with co-expression patterns
330 (dendrogram, Fig. 4A) revealing several distinct groupings. The "root" co-expression
331 group (orange box, Fig. 4A) includes three root samples and 12 DAP developing seed.
332 The "young inflorescence" co-expression group (blue box, Fig. 4A) includes male
333 (immature tassel) and female (immature cob) tissues. The "meristem" co-expression
334 group (green box, Fig. 4A) includes the developing embryo and growing shoot tip
335 tissues. These last two groups together form a larger co-expression cluster from
336 tissues that share the property of having actively dividing cell populations. Beyond
337 these three groups, several tissues show distinctive patterns of LINC gene expression,
338 not clustering with other tissues, including anthers (male pollen-bearing organs),
339 topmost seedling leaf, meiotic tassel, and silks (female, pollen-receiving organs).

340 To characterize the general expression levels of the candidate maize LINC
341 genes, we plotted the median FPKM values across different tissues (Fig. 4B). Maize
342 LINC genes generally exhibit low to moderate expression levels, similar to that of the

343 house-keeping *CDK1* gene (Fig. 4B). The more highly expressed LINC-related genes
344 had transcript levels that ranged from 80-140 median FPKM and included the
345 MKAKU4s, the C-terminal SUNs, the NCHs, and some of the KASH genes, *MLKP3* and
346 *MLKP4*. The lower expressed genes had transcript levels that ranged from 3-50
347 median FPKM and included the mid-SUNs, the MNEAPs, and most of the other KASH
348 genes. As a group, the maize AtWIP-like genes (*MLKP1-4*) have higher expression
349 levels than the AtSINE-like genes (*MLKS1*, *MLKS2*), with *MLKP3* and *MLKP4* showing
350 the most uniform and broad expression. In contrast, the mid-SUN *SUN5* gene exhibited
351 a most striking pattern of expression in only pollen. The *Poaceae*-specific gene,
352 *MLKG2,* appears to be expressed in tissues characterized by meristem and active cell
353 division, but absent from others such as endosperm, root, leaf, and silks.
354 Among the highly similar gene pairs that are likely to have arisen as part of a
355 recent genome duplication (Gaut et al., 2000), some of them (*MLKP1 and MKLP2;*
356 *MLKP3 and MLKP4; and MKAKU41 and MKAKU42*) show similar expression levels
357 whereas other gene pairs, such as *MNEAP1* and *MNEAP2* exhibit differential
358 expression levels. Overall, this expression analysis provides evidence for LINC and
359 associated proteins genes being transcribed and regulated in a tissue-specific manner.
360 The co-expression groupings predict possible but untested genetic or physical
361 interactions that may reflect shared functions.
362
363 **Identification of LINC and associated proteins from native co-IP using anti-**
364 **ZmSUN2 antibodies.**

365 In order to find native cellular interactors of SUN proteins, we carried out co-
366 immunoprecipitation (co-IP) optimized for maize NE proteins from purified earshoot
367 nuclei as summarized Table 2. Maize earshoots of ~3-5 cm in length yield high quality
368 cellular extracts from fast-growing somatic tissues several days prior to silking and
369 fertilization. The antibody used for native co-IP was the affinity-purified sera designated
370 anti-SUN2-NPAP as previously characterized (Murphy et al., 2014). Control co-IP
371 experiments were conducted using either no antibodies or irrelevant but similarly
372 affinity-purified rabbit sera, anti-ZmNDPK1 (Kopylov et al., 2015). A key technical
373 advance was the use of the short-chain crosslinker dithiobis(succinimidyl propionate),
374 DSP. We adapted the membrane protein cross-link immunoprecipitation (MCLIP)
375 method for use with maize because it was shown to effectively recover SUN-interacting
376 proteins from U2OS cells (Jafferali et al., 2014). We found that the MCLIP method
377 using DSP together with light formaldehyde fixation to preserve the nuclei resulted in
378 substantially increased final recovery and detection of known or predicted maize
379 homologs of plant LINC, NE, or SUN-interacting proteins (Tamura et al., 2013; Wang et
380 al., 2013; Zhou et al., 2012).
381 The proteins recovered from the co-IP assays were subjected to tryptic digestion
382 and identified by liquid chromatography tandem mass spectrometry (LC-MS/MS). Over
383 140 proteins were identified (Supplemental Table 2) using software thresholds of 95%
384 protein and 95% peptide. As expected, SUN2 was the most abundant protein, but we
385 also repeatedly detected SUN1, which could result from crossreactivity of ZmSUN1 with
386 our anti-SUN2 sera, or from actual co-IP of ZmSUN1 bound to SUN2, as expected from

387 documented hetero multimerization of SUN proteins in plants and animals (Sosa et al.,
388 2012).

389 The co-precipitated proteins identified were considered high confidence if they
390 were detected in multiple replicates but not in the negative controls. The specificity of
391 the modified MCLIP method is evidenced by recovery of core LINC components and
392 predicted cytoplasmic and nucleoplasmic proteins. Among the predicted core LINC
393 components repeatedly found in the top co-IP hits were three ONM proteins (MLKG1,
394 MLKS2, and MLKT2) two INM proteins (ZmSUN1 and ZmSUN2), two NMCP/CRWN
395 homologs (NCH1 and NCH2), two RAN family proteins (RAN GTPase and RAN-GAP2)
396 and five nuclear pore proteins (NUP133, NUP96, NUP88, GP210, and NUA).

397 We also found proteins that were co-precipitated but likely via indirect interaction
398 with SUN. These include the cytoskeletal proteins TUB6, TUB4 and Actin-1/ARP7. The
399 Actin-1/ARP7 protein could be cytoplasmic or nucleoplasmic or both because its
400 annotation includes the designation of nuclear actin, an intriguing possibility in light of
401 recent findings showing cross-species binding of Arabidopsis AtARP7 to the carrot
402 nuclear matrix protein DcNMCP1 (Mochizuki et al., 2017). We recovered several
403 nuclear proteins with DNA-binding activities including transcription factors bZIP, ARF10,
404 and the linker histone H1a. Overall, the co-IP resulted in confirmatory identification of
405 native SUN2-interacting proteins with known or suspected LINC proteins, plus several
406 additional proteins that represent new targets for future investigation. To further verify
407 some of the LINC candidates from either the Co-IP or bioinformatic analyses, we
408 selected three genes for further localization and interaction assays *in vivo* using
409 transient heterologous expression assays.

410

411 ***In planta* evidence of NE targeting and ZmSUN2 interaction for three candidate**

412 **maize LINC proteins**

413 In order to investigate the subcellular localization of the maize KASH proteins, we

414 transiently expressed N-terminal GFP fusions of MLKG1, MLKP1 or MLKP2 in *Nicotiana*

415 *benthamiana* leaf tissue as summarized in Figure 5. Expression of a positive control,

416 mCherry-ZmSUN2, shows that maize NE proteins can be properly expressed from the

417 plasmids and localized to the NE in this leaf cell heterologous expression system. All

418 three of the previously untested ONM maize LINC candidate proteins, MLKG1, MLKP1

419 and MLKP2, also localized to the nuclear periphery (Fig. 5A). These localization

420 patterns are comparable to those previously reported for AtKASH proteins in similarly

421 designed expression assays (Zhou et al., 2012; Zhou et al., 2014). For GFP-MLKG1, a

422 network-like pattern at the cell periphery was also seen (Fig. 5B). In order to further

423 examine this pattern, we co-expressed GFP-MLKG1 with an ER marker and an actin

424 marker and found that GFP-MLKG1 colocalizes with both the ER and the actin network

425 (Fig. 5B). Similar colocalization patterns have previously been demonstrated for

426 AtSINE1 (Zhou et al., 2014) but, unlike AtSINE1, MLKG1 does not possess an ARM

427 domain or other known actin binding motifs.

428 The C-termini of the KASH proteins are known to be required for NE localization

429 which is mediated through interaction with SUN proteins (Sosa et al., 2012; Zhou et al.,

430 2012; Zhou et al., 2014). To test whether this is also true of our maize KASH

431 candidates, we created and tested three classes of C-terminal deletions or replacement

432 constructs. ΔKASH represents constructs in which KASH domain, that is, all the amino

433 acid residues following the TMD are deleted (Fig. S1A). Δ4r, ΔKASHr and ΔTMr
434 represent constructs in which terminal four residues, KASH domain, or everything
435 beginning with TMD have been replaced with a non-genic 38 aa terminus from the
436 vector backbone, respectively (Fig. 5C, S1A).

437 We observed that each of these C-terminal deletion and replacement constructs
438 led to the loss of NE localization (Fig. 5C; Sup 1B). The modified proteins showed
439 localization patterns that resembled soluble proteins, visible throughout the cytoplasm
440 and nucleoplasm (Fig. 5C, supplemental figure 1). For comparison, we expressed GFP-
441 only as a control (Supplementary Figure 1B), which illustrates the distribution of a
442 soluble protein without a TM domain in the cytoplasm and nucleoplasm. Interestingly, in
443 every case, the Δ4 and Δ4r modifications were sufficient to abolish NE localization
444 despite the fact that the fusion protein still has a TMD. These results are consistent with
445 the idea that the terminal residues of these proteins may confer several properties
446 including stable post-translational membrane insertion and targeting to the ONM. In
447 order to investigate whether the NE localization seen for these maize KASH proteins
448 involves binding to maize SUN proteins, *in vivo*, we co-expressed three different maize
449 GFP-KASH with or without ZmSUN2 or deletion derivatives of ZmSUN2 (Figure 6A). If
450 they do interact, we would expect to see a reduction in the mobility of GFP-KASH as
451 measured by fluorescence recovery after photobleaching and indicative of interacting
452 partners, as previously described (Zhou et al., 2014). For two of the three, MLKG1 and
453 MLKP2, we did indeed observe reduced mobility when co-expressed with ZmSUN2 as
454 summarized in Fig. 6B. When GFP-MLKG1 is transiently expressed by itself, 63±8% of
455 the protein population is mobile while 36% remain immobilized. However, upon co-

456 expression with mCherry-ZmSUN2, only 49±12% of the protein population was mobile
457 indicating that the co-expressed ZmSUN2 substantially increased the immobile fraction
458 of GFP-MLKG1 from 36% to 51% (Fig. 6 A,B). We interpret this shift as indirect
459 evidence for ZmSUN2 interaction with GFP-MLKG1. To test this idea, GFP-MLKG1
460 was co-expressed with a domain deletion mutant of ZmSUN2 lacking the SUN domain
461 (mCherry-ZmSUN2ΔSUN) and found that the mobility of MLKG1 was similar to that of
462 the MLKG1 alone (68±7%; Fig. 6 B,C). This finding indicates that the SUN domain is
463 required for mediating the change in GFP-MLKG1 mobility, providing strong indirect
464 evidence that ZmSUN2 interacts with MLKG1. To further test if the coiled-coil domain of
465 the SUN protein is required for reduced mobility, we created a mCherry-ZmSUN2
466 lacking the coiled-coil domain, mCherry-ZmSUN2ΔCC. Co-expression of this mCherry-
467 ZmSUN2ΔCC with GFP-MLKG1 had a similar effect on KASH mobility to that of the full
468 length SUN2 (Fig. 6 B, C). This result establishes that the ZmSUN2-co-expression-
469 dependent increase in the immobile fraction of MLKG1 occurs to a similar degree with
470 or without the CC domain.
471 Similar changes in protein mobility were observed for GFP-MLKP2 (Fig. 6 F,G)
472 when co-expressed with full-length ZmSUN2, ZmSUN2ΔSUN or ZmSUN2ΔCC. Co-
473 expression with ZmSUN2 or ZmSUNΔCC resulted in significant increases in the
474 immobile fraction of GFP-KASH from 51% to 69% or 71%, respectively. to GFP-
475 MLKG1, co-expression of GFP-MLKP2 with ZmSUN2ΔSUN did not result in a
476 significant change in the mobile fraction of the maize KASH. In contrast, we did not
477 observe statistically significant changes in GFP-MLKP1 mobility when co-expressed
478 with full-length or deletion constructs of ZmSUN2 (Fig. 6 D, E), even though the

479 changes observed were all in the same directions as those for GFP-MLKG1 and GFP-
480 MLKP2. Representative FRAP images, individual repeat curve data and halftime
481 recovery values are shown in Supplementary Figures 2, 3, and 4.
482 Of the 3 candidates tested, MLKG1 was already found by native co-IP to interact
483 with SUN2, but the other two, MLKP1 and MLKP2 were not, presumably because of
484 their low expression levels in earshoot (Fig. 4). In order to check for their ability to
485 interact with SUN2 *in planta*, we assayed them by co-IP in tobacco tissues where they
486 were co-expressed as heterologous fluorescent fusion proteins (Supp. Fig 5A). Using
487 anti-GFP antisera, we observed co-precipitation of mCherry-SUN2, but not in control
488 extracts from non-infiltrated or mCherry-SUN2 only samples. These additional tests
489 confirmed that all three of the proteins tested, MLKG1, MLKP1, and MLKP2, showed
490 SUN2 interaction in one or more direct or indirect interaction assays (summary table in
491 Supp. 5B).
492 In summary, the fusion protein transient expression showed that all three
493 candidates tested showed localization to the NE, and deletion analysis established that
494 the last four amino acid residues were required for that localization. Furthermore,
495 these showed evidence of ZmSUN2 interaction *in planta* from one or more experiments,
496 inferred from co-expression FRAP or detected in native or heterologous co-IP
497 experiments. These findings provide the first experimental evidence for the
498 identification of multiple LINC complex components in monocot plant species,
499 evolutionarily distant from Arabidopsis and relevant to establishing the conservation of
500 the LINC complex across all eukaryotes.
501

502 # DISCUSSION

503 Here we set out to systematically identify all of the maize LINC core and
504 associated components as a way to advance nuclear envelope research. Despite the
505 wide variety of functionalities attributed to LINC complex in plants, animals, and yeast,
506 there remains a considerable number of unknowns regarding the mechanisms and
507 functions of individual components. Our findings represent a valuable model grass
508 crop species for future studies of the plant nuclear envelope and important biological
509 processes such as meiotic chromosome behavior, development, signaling, and
510 pathology (Kim et al., 2015; Kracklauer et al., 2013; Meier et al., 2017; Razafsky and
511 Hodzic, 2015; Tapley and Starr, 2013). Maize provides a valuable plant complement to
512 Arabidopsis, where SUN1 and SUN2 double knockouts result in disrupted meiotic
513 progression, excessive chromosomal interlocks, and reduced fertility (Varas et al.,
514 2015). However, Arabidopsis does not exhibit a canonical early prophase telomere
515 bouquet cluster on the nuclear envelope, nor has it been shown to have a meiotic SUN
516 belt (Murphy et al., 2014). These differences highlight one example of an important
517 principle that not all findings from one plant species can be generalized across the
518 kingdom. That, together with the difficulty in identifying orthologs illustrate the
519 importance of defining the composition of the LINC complex in multiple species where
520 genetic and cytological resources afford efficient and comparative analysis of the
521 associated biological functions.

522 Searching for maize KASH and associated ONM proteins, we found homologs of
523 other plant genes encoding AtWIP and AtSINE1/2-like proteins, but not those encoding
524 AtSINE3/4, SINE5 or AtTIK (Graumann et al., 2014; Zhou et al., 2014). Genes not found

525 in maize show limited evolutionary distributions. *AtSINE3/4* are found only in
526 brassicaceae, *SINE5* only in medicago, and *AtTIK* only to Arabidopsis (Poulet et al.,
527 2017a). Similarly, *MLKG1* and *MLKG2* genes are found only in maize or other grasses
528 but not eudicot species such as Arabidopsis or Medicago. In contrast, the *MLKT* genes
529 of maize are clearly homologous *AtWITs*, a rare case in which sequence conservation is
530 sufficient for identification of an ONM LINC component. It is worth noting, however, that
531 other maize genes resembling ONM LINC components were found but fell short of the
532 criteria we applied in this study. It remains an intriguing but recurring observation that
533 ONM LINC proteins contribute to ancient structure but often exist as highly divergent in
534 extant species. In this regard, plant KASH genes are similar to those from opisthokonts,
535 with characteristic variation in protein sequence and structure (Meier et al., 2017).
536 Searching for additional maize genes encoding INM proteins, we found three
537 homologs of the plant-specific *NEAP gene*s. The *NEAP* genes are limited to seed
538 plants (Poulet et al., 2017a) and not found in any of the representatives of the more
539 ancient moss and algal species included in this study (Fig. 3). The *MNEAP1* and
540 *MNEAP2* genes encode protein sequences that end in two basic residues, like their
541 eudicot homologs, but the functional significance of these terminal residues remains
542 unknown. Interestingly, the maize NEAPs have SMC_N domains which are currently
543 uncharacterized but may point to a chromosome tethering or nuclear structural role.
544 The maize NEAPs named in this study are three of the eleven maize genes assigned to
545 the PLAZA gene family HOM04M002962. Consequently, the actual number of the
546 maize NEAP genes may grow as revised bioinformatic and functional criteria for their

547 recognition emerges. In contrast, we did not find evidence for additional SUN genes of
548 maize beyond the five previously described (Murphy et al., 2010; Murphy et al., 2014).
549 Searching for nucleoplasmic proteins associated with plant LINC complexes, we
550 found two maize genes with homology to *NMCP*/*CRWN* genes, *NCH1* and *NCH2*, both
551 of which are deeply conserved in plants. Each of these are associated with multiple
552 transcript models (Wang et al., 2016), suggestive of splice variants, an intriguing
553 observation in light of the tendency for animal lamin genes to also have many splice
554 variants (Broers et al., 2006). Identification of plant nuclear matrix components that
555 interact with the nuclear envelope provides new opportunities to investigate the
556 mechanisms that underlie how signals are transduced via the LINC complex.
557 Cellular localization experiments confirmed the NE localization of fluorescent
558 fusion protein constructs for ZmSUN2 and all three of the maize KASH homologues
559 tested, MLKG1, MLKP1 and MLKP2 (Fig. 5). Interestingly, GFP-MLKG1 colocalizes
560 with the ER and the actin network, implicating it in cytoskeletal signal transduction to the
561 nucleus. Relatedly, Arabidopsis actin is tethered to the LINC complex through direct
562 interactions with AtSINE1 and through myosin XI-i - WIT-WIP interactions (Tamura et
563 al., 2013; Xu et al., 2007; Zhou et al., 2014). The presence of *SINE1* (*MLKS1* and
564 *MLKS2*), *WIP* (*MLKP1* - *MLKP4*) and *WIT* (*MLKT1* and *MLKT2*) homologs in maize,
565 together with the poaceae-specific *MLKG1* suggests that plant genomes may encode a
566 diverse repertoire of of proteins to connect and integrate NE with cytoskeletal
567 processes.
568 To provide additional experimental evidence on the composition of the maize
569 LINC complex, we used a ZmSUN2 Co-IP method optimized for membrane proteins

570 together with isolated nuclei. We found a large number of different proteins that were
571 generally consistent with predictions from our working model (Fig. 1) and previous
572 reports (Dittmer and Richards, 2008; Goto et al., 2014; Pawar et al., 2016; Tamura et
573 al., 2010; Tamura et al., 2013; Wang et al., 2013; Zhou et al., 2012; Zhou et al., 2015b;
574 Zhou et al., 2015c). Among the expected and reported LINC and LINC-associated
575 proteins, our native Co-IP detected SUN, WIPs, WIT, NCH, and RAN cycle proteins.
576 In relation to the cytoplasmic and ONM side of the LINC complex, we detected
577 proteins predicted to bind directly to ZmSUN2 including MLKG1 and MLKS2. We also
578 co-precipitated proteins predicted to bind indirectly to ZmSUN2 including MLKT2, two
579 RAN cycle proteins, RAN GTPase, RAN-GAP2, Tub4, Tub6, and the nuclear pore
580 proteins NUP133, NUP96, NUP88, GP210, and NUA. Notably, we did not detect all of
581 the nuclear pore proteins, consistent with previous studies showing interactions of a
582 specific and limited subset of NPC proteins with SUN-domain proteins in association
583 with transcriptional activation and mRNA export processes (Gerber et al., 2004; Li and
584 Noegel, 2015). In relation to the nucleoplasmic and INM side of the LINC complex, we
585 detected two INM proteins, ZmSUN1 and ZmSUN2; and two NMCP/CRWN proteins,
586 NCH1 and NCH2.
587 The Co-IP experiments also yielded an interesting list of additional proteins
588 beyond the core LINC and associated proteins summarized in our working model.
589 Some of these proteins appear to belong to non-specific protein categories, but in fact
590 have been found in various NE studies including nuclear actin (ARP7/Actin-1),
591 transcription factors (bZIP5, ARF10, MYBR35, bHLH72/bHLH041), Histone H1 (H1a),
592 and ribosomal proteins (L19, L6, and L22-2) (Horigome et al., 2011; Hotta et al., 2007;

593 Mochizuki et al., 2017; Nakayama et al., 2008). Nuclear actin is increasingly recognized
594 as playing key roles in chromosome dynamics and gene expression (reviewed by
595 (Spichal and Fabre, 2017) whereas a plant linker H1 histone was recently shown to
596 have MTOC activity and reside at the nuclear periphery (Boruc et al., 2012; Nakayama
597 et al., 2008).

598 Given the increasing recognition that the NE and nucleoskeletal proteins are
599 implicated in chromatin remodeling and gene regulation, these additional, non-LINC
600 proteins discovered through native co-IP represent ideal targets and experimental
601 inroads for mechanistic studies of LINC-mediated dynamic processes coupled to
602 chromatin structure and nuclear architecture. For instance, the NE cooperates with
603 nucleoskeletal proteins to modulate euchromatin and heterochromatin organization
604 (Poulet et al., 2017b; Wang et al., 2013). Emerging breakthroughs in HiC mapping of
605 nuclear architecture are likely to uncover additional principles of genome organization
606 that involve LINC-binding nuclear proteins (Grob et al., 2014). Given the importance of
607 lamins in human health, it is possible that plant NMCP/CRWN/NCH defects may even
608 model animal disease processes.

609 Both Co-IP and FRAP provide complementary assays for investigating *in planta*
610 binding of maize KASH candidates to ZmSUN2 (Fig. 6). For example, MLKG1 was co-
611 purified from maize extracts using SUN2 antisera (Table 2) while FRAP assays
612 demonstrated a ZmSUN2-dependent change in the mobility of GFP-MLKG1 and GFP-
613 MLKP2. The heterologous expression system described in this and previous studies
614 (Zhou et al., 2012; Zhou et al., 2014) demonstrates an efficient means to detect,
615 isolate, and dissect protein-protein interactions for distantly related grass species such

616 as maize, which is not easily transformed. They also provide an efficient means to carry
617 out *in vivo* tests of candidate interactors (e.g. Supp Fig 5) for which biochemical
618 amounts of native tissue may be limiting.
619 Overall, this study provides a comprehensive list of maize LINC candidates (Fig.
620 1) with multiple lines supporting experimental evidence. Through our efforts to identify
621 the the full ensembl of LINC components in a model eukaryote, we have contributed to
622 valuable new avenues of research to address unanswered questions regarding the role
623 of the nuclear envelope as an integrative structure coordinating nuclear and cellular
624 functions.
625

626 ## METHODS

627

628 **Bioinformatic identification of MLKs**

629 The maize B73 proteome v4
630 (ftp://ftp.gramene.org/pub/gramene/CURRENT_RELEASE/fasta/zea_mays/pep/) was
631 searched for sequences ending in motif [DTVAMPLIFY][VAPIL]PT (brackets indicate
632 alternative amino acid residues at the respective position). The selected proteins were
633 further searched for the presence of TMD using TMHMM (Krogh et al., 2001) and
634 Phobius softwares (Käll et al., 2004). Proteins lacking TMD or specific ending were
635 discarded. The selected maize MLKs were then searched on PLAZA monocot v4
636 (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_monocots/) and those
637 belonging to same homolog family were grouped together.
638

639 **Protein feature calculations and domain searches**

640 Theoretical Molecular weight and isoelectric point (pI) of candidate proteins was
641 calculated using pI/Mw tool of ExPASy (https://web.expasy.org/compute_pi/). Coiled coil
642 domains were identified by Jpred (http://www.compbio.dundee.ac.uk/jpred/). For the
643 identification of conserved domain sequences, the candidate protein sequences were
644 searched on Uniprot (http://www.uniprot.org/) , NCBI Conserved Domain Database
645 (https://www.ncbi.nlm.nih.gov/cdd/), PFAM (https://pfam.xfam.org/), SMART
646 (http://smart.embl-heidelberg.de/). The domains identified were then plotted using IBS
647 1.0.3 (Illustrator for Biological sequences) software (Liu et al., 2015).

648

649 **Identification of LINC orthologs in other organisms and phylogenetic**
650 **reconstruction**

651 Sequences of orthologs of Maize LINC candidates belonging to matching HOMO gene
652 families were retrieved from PLAZA v4
653 (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_monocots/). Sequences of
654 all species selected for this study are present in PLAZA except *Medicago truncatula*,
655 which were retrieved from Gramene. The protein sequences were further screened for
656 the presence of characteristic protein features. The MLKS/MLKG/MLKP orthologs were
657 searched for the presence of TMD towards C-terminus and for terminal four residues of
658 the proteins to be VIPT, VVPT, AVPT, PLPT, TVPT, LVPT or PPPS. For MLKT
659 orthologs, proteins having high sequence similarity and C-terminal TMD were retained.
660 For NCHs, KAKUs and NEAPs, sequences having high sequence similarity to
661 Arabidopsis CRWNS, KAKU4 and NEAPs were selected. For phylogenetic

662 reconstruction of each protein category, sequences were first aligned using MUSCLE
663 and Neighbour end joining trees were generated in MEGA 7.0.21 (Kumar et al., 2016)
664 using default parameters.
665
666 **Transcriptome**
667 The expanded maize gene expression atlas (Stelpflug et al., 2016) was
668 downloaded for analysis. Log_{10} Fragments Per Kilobase of exon per Million fragments
669 mapped (FPKM) values for 15 tissues were used to construct heatmap using Heatmap2
670 package of R. To group tissues by similar expression profiles, hierarchical clustering
671 was done using Hclust function of R.
672
673 **Co-IP protocol**
674 Maize (Zea mays L) B73 seeds were obtained from WF Thompson (NC State
675 University, Raleigh, NC, USA, accession “27 FARMS”, bulked for NCSU, denoted
676 B73NC27). Seeds were grown at the Florida State University Mission Road Research
677 Facility (Tallahassee, FL). Young earshoots (3-5 cm long) were harvested from 7-8-
678 week-old plants, flash frozen in liquid nitrogen and stored at -80 °C until use. Ten
679 grams tissue, were ground under liquid nitrogen with a mortar and pestle and cross-
680 linked with 0.1% formaldehyde by stirring for 10 min in 100 mL ice-cold Buffer A (15
681 mM PIPES-NaOH, pH 6.8, 0.32 mM sorbitol, 80 mM KCl, 20 mM NaCl, 0.5 mM EGTA,
682 2 mM EDTA, 1 mM DTT, 0.15 mM spermine, and 0.5 mM spermidine; (Belmont et al.,
683 1987). The nuclei were isolated as described in (Vera et al., 2014) and resuspended in
684 5 ml PBS. The isolated nuclei were treated with 2 mM DSP (dithiobis(succinimidyl

685 propionate) (Cat. No. 22585, Thermo Scientific) crosslinker for 2 h at 4 °C. To lyse the
686 nuclei, 7 M urea and 1% Triton X-114 and Plant protease inhibitor cocktail (Cat. No.
687 P9599, Sigma) were added and incubated for 20 min on ice. The ruptured nuclei were
688 sonicated three times in Bioruptor (Diagenode) for 5 minutes each at medium setting.
689 Urea was diluted with 8 volumes of PBS. Homogenates were centrifuged at 10,000g
690 for 10 min at 0 °C to remove cellular debris. The supernatant was collected and pre-
691 cleared by adding 0.5 mg Protein A dynabeads (Cat No.10000D, Thermo Fisher) for 30
692 min with rotation at 4 °C. The beads were removed by placing tubes on magnetic rack
693 (Cat. No.12321D, Thermo Fisher). The cleared lysate was split into 3 equal aliquots. 3
694 µg α- SUN2 antibody was added to first aliquot, 3 µg NDPK control antibody to the
695 second tube and no antibody added to the third tube. The tubes were incubated at 4 °C
696 with rotation for overnight followed by the addition of 0.5 mg Protein A magnetic beads
697 to each tube for one-hour incubation at 4 °C. The antigen-antibody bound beads were
698 captured with the help of magnetic rack and washed thrice with PBS supplemented
699 with 0.02% Tween-20. The protein complexes bound to the beads were digested on
700 the using ProteoExtract All-in-One Trypsin Digestion Kit (Cat. No. 650212,
701 Calbiochem) according to manufacturer's instructions. Briefly, beads were
702 resuspended in Extraction buffer and digestion buffer was added, reduced using the
703 provided reducing agent for 10 minutes at 37°C. Samples were cooled to RT and then
704 blocked using blocking reagent for 10 minutes at room temperature. Trypsin at a final
705 concentration of 8 ng/µl was added and incubated for 2 hours at 37°C with shaking.
706 The reaction was stopped by acidification. Digests were vacuum dried and submitted
707 for LC-MS/MS to Translational Lab at Florida State University.

708

709 **Mass spectrometry and data analysis**

710 Digested peptides were resuspended in 0.1% Formic acid for loading onto LC-MS at
711 Translational Lab, Florida State University. An externally calibrated Thermo LTQ
712 Orbitrap Velos nLC-ESI-LTQ-Orbitrap (high-resolution electrospray tandem mass
713 spectrometer) was used with the following parameters. A 2 cm, 100 µm i.d. (internal
714 diameter) trap column (SC001 Easy Column from Thermo-scientific) was followed by a
715 10 cm analytical column of 75 µm i.d. (SC200 Easy Column from Thermo-scientific).
716 Both trap column and analytical column had C18-AQ packaging. Separation was carried
717 out using Easy nanoLC II (Thermo-Scientific) with a continuous, vented column
718 configuration. A 5 µL (~500 ng) sample was aspirated into a 20 µL sample loop and
719 loaded onto the trap. The flow rate was set to 300 nL/min for separation on the
720 analytical column. Mobile phase A was composed of 99.9 H_2O (EMD Omni Solvent),
721 and 0.1% formic acid and mobile phase B was composed of 99.9% ACN, and 0.1%
722 formic acid. A 1 h linear gradient from 0% to 45% B was performed. The LC eluent was
723 directly nano-sprayed into an LTQ Orbitrap Velos mass spectrometer (Thermo
724 Scientific). During the chromatographic separation, the LTQ Orbitrap Velos was
725 operated in a data-dependent mode and under direct control of the Xcalibur software
726 (Thermo Scientific). The MS data were acquired using the following parameters: 10
727 data-dependent collisional-induced-dissociation (CID) MS/MS scans per full scan. All
728 measurements were performed at room temperature and three technical replicates. The
729 raw files were analyzed using Proteome Discoverer (version 1.4) software package with
730 SequestHT and Mascot search nodes using maize proteome database available at

731 Gramene (AGVP3.0). The resulting msf files were further analyzed by the 'Scaffold
732 version 4.4' the proteome validator software at 90% peptide and protein threshold with a
733 minimum of 1 peptide.
734
735 **Constructs and gene cloning**
736 Coding sequences of MLKP1 and MLKP2 genes fused to C-terminus of eGFP-FLAG-
737 HA sequence were custom synthesized by Genscript as pHWBF08 and pHWBF06
738 respectively. eGFP-FLAG-HA-MLKP2 and eGFP-FLAG-HA-MLKP2 sequences were
739 PCR amplified with att flanking primers and cloned individually into pDONR221 vector
740 by BP cloning (Cat. No. 1235019, Invitrogen) to generate pHWBF08EC and
741 pHWBF06EC entry clones respectively. These were moved to destination vector
742 pH7WG2 (Karimi et al., 2002) by LR recombination reaction of Gateway cloning system
743 (Cat. No. 1235019, Invitrogen) to obtain expression vectors pPK1iexp and pPK1Cexp
744 respectively. For MLKG1, coding region was PCR amplified from maize B73 using gene
745 specific primers (ST3) and cloned into pCRII vector by TOPO cloning (Cat. No.
746 K461020, Thermo Fisher). After sequence verification, the coding sequence was
747 amplified with forward primer flanked with BamHI restriction site sequence and reverse
748 primer flanked by SbfI restriction site sequence. The PCR amplified product was double
749 digested with BamHI and SbfI (NEB) and ligated in pECGFP (Bass lab) donor vector in
750 frame and downstream of eGFP-FLAG-HA sequence. The fusion sequence was then
751 moved to pH7WG2 destination vector by LR recombination to generate expression
752 vector called pPK1Dexp. Deletion constructs for MLKP1 and MLKP2 were generated
753 using corresponding full length entry clones as template to PCR amplify required

754 regions with internal primers containing flanking att sequences (ST3) without a stop
755 codon and cloned into pDONR211 by BP cloning. These deletion constructs were
756 subcloned in pH7WG2 destination vector by LR recombination and the final expression
757 cassettes contains additional 38 aa residues
758 "NPAFLYKVVISRGHARVRKNHQSLSTNLSLSIFLQNNV" from the non-genic region of
759 the vector backbone at their C-termini.
760 MLKG1ΔKASH was generated using Q5 DNA mutagenesis kit (Cat. No. E0554S,
761 NEB). Briefly, two primers were designed in opposite directions, going outwards from
762 KASH domain. The entire entry vector (pECPK1D) minus the KASH domain was
763 amplified using high fidelity Q5 polymerase provided with the kit. The template
764 contamination was removed and vector re-ligated using manufacturer provided enzyme
765 mix containing a kinase, a ligase and DpnI to create entry clone pPK1DdelKASH. The
766 entry clone were then moved to pH7WG2 destination vector by LR recombination.
767
768 ***Agrobacterium tumefaciens* transformation**
769 *A. tumefaciens* was transformed with expression vectors using established heat shock
770 protocols (Graumann and Evans, 2010). Briefly, competent cells were thawed on ice
771 and 4ul plasmid DNA was added. Cells were incubated for 5 min on ice, 5 min in liquid
772 nitrogen and 5 min at 37 °C 1 ml LB medium was added and cells were incubated for 2h
773 at 28 °C before being plated out and left to grow for 48 h at 28 °C.
774
775 ***Nicotiana benthamiana* transformation and microscopy**

776 Transient expression of fluorescent fusion proteins was carried out as previously
777 described (Sparkes et al., 2006). Briefly, Agrobacteria were suspended in infiltration
778 buffer at an OD of 0.05 and pressed into the lower leaf epidermis with a syringe. 3 days
779 post infiltration, leaf sections were imaged with a Zeiss LSM 800 confocal microscope
780 using a 100X oil immersion objective, NA 1.46 and 1.5X digital zoom. GFP was excited
781 at 488 nm and emission collected between 500-550 nm, mCherry was excited with a
782 561 nm laser and emission collected from 565-620 nm. For FRAP, 5 pre scans and then
783 a 160x160 pixel ROI is bleached with the 488 nm laser at 100% for 10 iterations.
784 Recovery data was collected over 36 s (240 scans, 0.13 s framerate). Mean ROI
785 intensity data was extracted using Zen Blue (version 2.3) and normalised as described
786 in (Martinière et al., 2012). One phase association recovery curves, plateau values and
787 halftimes were produced in Graphpad (version 6). ANOVA statistical test with multiple
788 comparisons was performed. For DAPI staining, leaf sections were first incubated in 1
789 ml water containing 5 ul 20% Triton X-100 and 1ul DAPI. DAPI excitation was
790 performed with the 405 nm laser and emission between 400-450 nm imaged. DAPI/GFP
791 or DAPI/mCherry imaging was performed using multitrack line scanning.
792
793 **Direct Co-IP and western blot**
794 1 g of *N. benthamiana* leaves were harvested and ground in liquid nitrogen.
795 The powdered tissue was homogenized in 5 mL of radioimmunoprecipitation assay
796 buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% SDS, 0.5% Na-deoxycholate, 1%
797 NP-40, 1 mM PMSF, and 1% protease inhibitor cocktail (Cat. No. P9599, Sigma) in a
798 dounce homogenizer kept on ice. The homogenate was centrifuged at 10,000 xg and

799 the supernatant was collected for co-IP. One tenth of the protein extracts was used as
800 the input sample and concentrated using acetone precipitation. The rest of the lysate
801 was subjected to co-IP using dynabeads coated with rabbit anti-GFP antibody (Cat. No.
802 ab290, Abcam). The lysate was incubated with beads for 4 hours after which the beads
803 were collected using magnetic rack and washed with PBST three times. The antibody-
804 protein complex bound to beads was eluted using 50 mM Glycine (pH 2.8). The
805 immunoprecipitates and the input samples were boiled in 4X Laemelli's gel loading
806 buffer, fractionated on 15% SDS-PAGE and transferred to polyvinylidene difluoride
807 membranes (Cat. No. 88518. Thermo Scientific), and detected with a mouse anti-GFP
808 (1:3,000; Cat No. MA5-15256,Thermo Scientific) or rabbit anti-SUN2-NPAP (1:1,000;
809 Murphy et al 2014) antibody at room temperature for 1 h. After three 15-min washes in
810 PBS-T buffer at room temperature, the membranes were incubated with a 1:3,000
811 dilution of anti-rabbit or anti-mouse IgG horseradish peroxidase–linked antibody (Santa
812 Cruz Biotechnology, Santa Cruz, CA) for 1 h at room temperature, then subjected to 3 ×
813 15 min washes in PBST buffer at room temperature. The PVDF membrane was
814 incubated with chemiluminescent reaction kit for 3-5 min at room temperature (Millipore,
815 Immobilon detection kit, WBKL50100, Billerica, MA) and exposed to X-ray film in dark
816 room to visualize the protein bands.
817
818

819 ## FIGURE LEGENDS

820 **Figure 1**. **Working model of the plant LINC complex; core and associated**
821 **proteins.** Core LINC and associated proteins are depicted as residing in one of the four

822 cellular locations, each represented by different background shaded colors: green,
823 cytoplasm (CYTO); yellow, outer nuclear membrane (ONM); pink inner nuclear
824 membrane (INM); and blue NUC (nucleoplasm). The core LINC proteins (dashed box
825 enclosed) shown include SUN-domain proteins labeled "SUN", KASH-related proteins
826 "KASH/MLK", and AtWIT-related proteins labeled "WIT/MLKT". The LINC-associated
827 proteins shown include cytoplasmic proteins labeled "RanGAP" and "motor proteins",
828 INM proteins labeled "NEAP/MNEAP", and nucleoplasmic proteins labeled
829 "NMCP/CRWN/NCH" and "KAKU4/MKAKU4". Maize gene names and their known
830 Arabidopsis homolog types are listed. This composite model places proteins together on
831 the basis of previously documented pair-wise interactions representing a generalized
832 consensus assembly but not meant to depict a single complex.

833

834 **Figure 2. Domain organization of candidates for the maize core LINC and**
835 **associated proteins complex.** Schematic representations of predicted proteins from
836 canonical transcript models are diagrammed to show the location of domains (color
837 coded and listed at bottom). The proteins are drawn to scale, using a separate scale for
838 the last two larger proteins, NCH1 and NCH2. The four C-terminal residues and stop
839 codon (*) location of the KASH and MNEAP group proteins are indicated. The
840 previously published SUN-domain proteins (Murphy et al., 2010) are not included. The
841 partial SMC N-terminal domains in the MNEAPs are not diagrammed but are located at
842 AA 24-308 for MNEAP1, AA 18-309 for MNEAP2, and at AA 3-109 for MNEAP3.
843 Potential CDK1 phosphorylation sites are shown (blue circles) for NCH1 and NCH2.

844

845 **Figure 3. Evolutionarily conservation of plant LINC complex components**.
846 Candidate maize LINC gene names, cellular location predictions, gene family group,
847 and PLAZA gene family ID (each starting with shared prefix of HOMO4M) are listed in
848 the first four columns. The other columns provide counts of the numbers of genes for
849 each species (top row) in each Plaza gene family (4th column). The general grouping
850 color key is indicated on the top left of the figure. Species included are Zema, *Zea*
851 *mays;* Sobi*, Sorghum bicolor,* Seit, *Setaria viridis;* Hovu, *Hordeum vulgare;*
852 Trae,*Triticum aestivum;* Brdi, *Brachypodium distachyon;* Orsa*, Oryza sativa spp.*
853 *Japonica;* Zoma*, Zostera marina;* Arth*, Arabidopsis thaliana;* Potr*, Populus trichocarpa;*
854 Vivi*, Vitis vinifera;* Metr*, Medicago truncatula;* Piab*, Picea abies;* Phpa*, Physcomitrella*
855 *patens* and Chre, *Chlamydomonas reinhardtii.* Broader grouping names are provided
856 along the top (labeled grey boxes).

857

858 **Figure 4: Gene expression levels and co-expression clustering of 20 maize LINC**
859 **genes.** Expression levels of 20 maize candidate LINC genes were analyzed using
860 published transcriptome data (Stelpflug et al., 2016) and including the genes
861 designated as housekeeping genes (HKG) in a previous study (Lin et al., 2014). (A)
862 Expression profiles of LINC genes in several tissues developmental stages. Data are
863 presented as log10 Fragments Per Kilobase of exon per Million fragments mapped
864 (FPKM) values from published transcriptome. Hierarchical clustering grouped tissues
865 according to similar expression profiles. The root, young inflorescence, and meristem
866 groups are denoted by orange, blue, and green boxes, respectively. The growth stage
867 of maize plant at the time of tissue harvesting are shown in parentheses; V, vegetative

868 stage; R, reproductive stage; DAP, days after pollination. (B) The Median FPKM values
869 of the tissues shown in panel A are plotted, ranked by value for individual genes.
870 Housekeeping genes are shown in red.

871

872 **Figure 5. Subcellular localisation of transiently expressed GFP-MLKP fusions in**
873 ***N. benthamiana*.** Agrobacterium-mediated leaf transformation with expression
874 constructs was carried out followed by imaging three days later using 100X lense with
875 confocal microscopy to detect DAPI (405 nm excitation laser), GFP (488 nm excitation
876 laser), or mCherry (561 nm excitation laser). Representative focal sections are shown
877 for GFP (green) or mCherry (purple). (A) GFP-MLKG1, GFP-MLKP1, GFP-MLKP2 and
878 mCherry-ZmSUN2 (green) are localised at the nuclear periphery, surrounding the
879 nucleus (n) stained with DAPI (magenta). (B**)** GFP-MLKG1 (green) is also present in a
880 network-like pattern at the cell periphery, where it partially co-localises with the ER
881 marker HDEL-mRFP (magenta, top row) and the actin marker lifeact-mRFP (magenta,
882 bottom row). (C**)** Schematic overview depicting the terminal deletions that were replaced
883 with 38 aa extension, $(N)_{38}$. (D) GFP-MLKP2Δ4r, GFP-MLKP2ΔKASHr and GFP-
884 MLKP2Δ4TMr (green) localise in the cytoplasm and the DAPI-stained nucleoplasm
885 (magenta). All scale bars 10µm.

886

887 **Figure 6. Fluorescence recovery after photobleaching (FRAP) co-expression**
888 **interaction assays of MLK LINC candidates with ZmSUN2 in *N. benthamiana*.** (A)
889 Schematic overview depicting the full length mCherry-ZmSUN2 and internal deletion

890 constructs that removed the coiled coil domain (ΔCC) or the SUN domain (ΔSUN).
891 Normalized averaged (n=26-37) intensity FRAP curves for (B) GFP-MLKG1, (D) GFP-
892 MKLP1, or (F) GFP-MLKP2 expressed alone (green), with full-length mCherry-SUN2
893 (black), with mCherry-SUN2-ΔCC (blue), or with mCherry-SUN2-ΔSUN (pink). Plateau
894 values are plotted to the right for assays with GFP-MLKG1 (C), GFP-MLKP1 (E), and
895 GFP-MLKP2 (G). Box plot error bars represent SD (blue) and mean (red) values.
896 ANOVA statistical analyses with multiple comparisons are noted (ns = not significant =
897 p≥0.05, **** = significant with P ≤0.0001).

898

899 **Supplementary Figure 1. Subcellular localisation of transiently expressed GFP-**
900 **MLKp domain deletion fusions in *N. benthamiana*.** Full-length, domain deletions or
901 replacements, and GFP only controls were assayed as described in Figure 5. (A)
902 Schematic overview of domain composition of the constructs. Constructs are listed at
903 the left of each row for images take through the middle of the nuclei (n in DAPI,
904 magenta). The right-most images (Cell periphery) were taken at a different focal plane
905 in order to image the cytoplasm near the cell cortex. Free GFP (bottom row, green) is
906 observed in the cytoplasm and, to a lesser extent, in the DAPI-stained nucleoplasm (
907 magenta). All scale bars 10μm.

908

909 **Supplementary Figure 2. FRAP images from the co-expression interaction**
910 **assays.** Representative images of cells during the FRAP assays for the experiments
911 summarized in Figure 6. The FRAP ROIs (white boxes) are shown along with images

912 taken at the times indicated across the top for (A) GFP-MLKG1, (B) GFP-MLKP1, and
913 (C) GFP-MLKP2. Scale bar denotes 2µm. FRAP ROI shown by white box. All scale
914 bars 2µm.

915

916 **Supplementary Figure 3. Normalised intensity FRAP curves co-plotted from all**
917 **experiments performed in *N. benthamiana* leaves.** The GFP fusion protein
918 constructs (MLKG1, MKLP1, and MLKP2) are grouped by columns as labeled at top.
919 The co-expression construct are grouped by row as labeled at left. Individual data
920 points (blue dots), single experiment recovery curves (black lines) and global average
921 recovery curves (red lines) and replicate numbers (n) are shown in each panel.

922

923 **Supplementary Figure 4.** FRAP recovery halftime values. The recovery half times
924 (t1/2(s)) are plotted for the full-length and co-expression experiments summarized in
925 Figure 6. The SD (blue lines) and mean (red lines) are shown for (A) GFP-MLKG1, (B)
926 GFP-MLKP1, and (C) GFP-MLKP2. ANOVA statistical analyses with multiple
927 comparisons are noted (ns = not significant = $p \geq 0.05$, ** = significant with $P \leq 0.01$, ****
928 = significant with $P \leq 0.0001$). Replicate numbers range from n=26 to n=35, as listed in
929 Supplementary Figure 3.

930

931 **Supplementary Figure 5. Co-immunoprecipitation assays of MLKP1 and MLKP2**
932 **with ZmSUN2 in *N. benthamiana*.** Total protein extracts from *N. benthamiana* leaves

933 harvested 3 days post-infection were subjected to co-IP using dynabeads coated with
934 anti-GFP antibody. (A) Western Blot of the co-precipitates and inputs with primary
935 antibodies listed on the right side of the blot. The FP construct combinations used for
936 infiltration are listed on the top of the blot. The numbers on the left are protein molecular
937 weight standards. (B) A summary of KASH-SUN interaction evidence tabulated from
938 this experiment (Heterologous co-IP) or other experiments in this study (Native co-IP
939 from Table 2, Heterologous FRAP form Figure 6 and Supplemental Figures 2-4).

940

941 ## ACKNOWLEDGEMENTS

942 Thank Eric Richards and Iris Meier for helpful and insightful comments on the
943 manuscript, Allison Jevitt for artwork on Figure 1, Dr. Rakesh K Singh and staff of the
944 Translational lab (College of Medicine, FSU) for assistance and advice on Co-IP and
945 Mass spectrometry. The work was supported by grants from the National Science
946 Foundation to HWB (NSF IOS 1444532) and Oxford Brookes University to KG, and
947 from a Dissertation Research Grant to HKG (Florida State University Graduate School).

948 ## Competing interests

949 The authors declare no competing or financial interests

950

951 ## REFERENCES

952 **Belmont, A. S., Sedat, J. W. and Agard, D. A.** (1987). A three-dimensional approach
953 to mitotic chromosome structure: evidence for a complex hierarchical organization.
954 *J. Cell Biol.* **105**, 77–92.

955 **Boruc, J., Zhou, X. and Meier, I.** (2012). Dynamics of the plant nuclear envelope and
956 nuclear pore. *Plant Physiol.* **158**, 78–86.

957 **Broers, J. L. V., Ramaekers, F. C. S., Bonne, G., Yaou, R. B. and Hutchison, C. J.**
958 (2006). Nuclear lamins: laminopathies and their role in premature ageing. *Physiol.*
959 *Rev.* **86**, 967–1008.

960 **Burke, B. and Stewart, C. L.** (2014). Chapter One - Functional Architecture of the
961 Cell's Nucleus in Development, Aging, and Disease. In *Current Topics in*
962 *Developmental Biology* (ed. Stewart, C. L.), pp. 1–52. Academic Press.

963 **Chettoor, A. M., Givan, S. A., Cole, R. A., Coker, C. T., Unger-Wallace, E.,**
964 **Vejlupkova, Z., Vollbrecht, E., Fowler, J. E. and Evans, M. M.** (2014). Discovery
965 of novel transcripts and gametophytic functions via RNA-seq analysis of maize
966 gametophytic transcriptomes. *Genome Biol.* **15**, 414.

967 **Ciska, M., Masuda, K. and Moreno Díaz de la Espina, S.** (2013). Lamin-like
968 analogues in plants: the characterization of NMCP1 in Allium cepa. *J. Exp. Bot.* **64**,
969 1553–1564.

970 **Coates, J. C.** (2003). Armadillo repeat proteins: beyond the animal kingdom. *Trends*
971 *Cell Biol.* **13**, 463–471.

972 **Crisp, M., Liu, Q., Roux, K., Rattner, J. B., Shanahan, C., Burke, B., Stahl, P. D. and**
973 **Hodzic, D.** (2006). Coupling of the nucleus and cytoplasm: role of the LINC
974 complex. *J. Cell Biol.* **172**, 41–53.

975 **Dittmer, T. A. and Richards, E. J.** (2008). Role of LINC proteins in plant nuclear
976 morphology. *Plant Signal. Behav.* **3**, 485–487.

977 **Fridkin, A., Penkner, A., Jantsch, V. and Gruenbaum, Y.** (2009). SUN-domain and
978 KASH-domain proteins during development, meiosis and disease. *Cell. Mol. Life*
979 *Sci.* **66**, 1518–1533.

980 **Friederichs, J. M., Ghosh, S., Smoyer, C. J., McCroskey, S., Miller, B. D., Weaver,**
981 **K. J., Delventhal, K. M., Unruh, J., Slaughter, B. D. and Jaspersen, S. L.** (2011).

982 The SUN protein Mps3 is required for spindle pole body insertion into the nuclear
983 membrane and nuclear envelope homeostasis. *PLoS Genet.* **7**, e1002365.

984 **Gaut, B. S., Le Thierry d'Ennequin, M., Peek, A. S. and Sawkins, M. C.** (2000).
985 Maize as a model for the evolution of plant nuclear genomes. *Proc. Natl. Acad. Sci.*
986 *U. S. A.* **97**, 7008–7015.

987 **Gerber, A. P., Herschlag, D. and Brown, P. O.** (2004). Extensive association of
988 functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in
989 yeast. *PLoS Biol.* **2**, E79.

990 **Goto, C., Tamura, K., Fukao, Y., Shimada, T. and Hara-Nishimura, I.** (2014). The
991 Novel Nuclear Envelope Protein KAKU4 Modulates Nuclear Morphology in
992 Arabidopsis. *Plant Cell* **26**, 2143–2155.

993 **Graumann, K. and Evans, D. E.** (2010). Plant SUN domain proteins: components of
994 putative plant LINC complexes? *Plant Signal. Behav.* **5**, 154–156.

995 **Graumann, K., Runions, J. and Evans, D. E.** (2010). Characterization of SUN-domain
996 proteins at the higher plant nuclear envelope. *Plant J.* **61**, 134–144.

997 **Graumann, K., Vanrobays, E., Tutois, S., Probst, A. V., Evans, D. E. and Tatout, C.**
998 (2014). Characterization of two distinct subfamilies of SUN-domain proteins in
999 Arabidopsis and their interactions with the novel KASH-domain protein AtTIK. *J.*
1000 *Exp. Bot.* **65**, 6499–6512.

1001 **Grob, S., Schmid, M. W. and Grossniklaus, U.** (2014). Hi-C analysis in Arabidopsis
1002 identifies the KNOT, a structure with similarities to the flamenco locus of Drosophila.
1003 *Mol. Cell* **55**, 678–693.

1004 **Groves, N. R., Biel, A. M., Newman-Griffis, A. H. and Meier, I.** (2018). Dynamic
1005 Changes in Plant Nuclear Organization in Response to Environmental and
1006 Developmental Signals. *Plant Physiol.* **176**, 230–241.

1007 **Hagan, I. and Yanagida, M.** (1995). The product of the spindle formation gene sad1+
1008 associates with the fission yeast spindle pole body and is essential for viability. *J.*

1009 *Cell Biol.* **129**, 1033–1047.

1010 **Haque, F., Lloyd, D. J., Smallwood, D. T., Dent, C. L., Shanahan, C. M., Fry, A. M.,**
1011 **Trembath, R. C. and Shackleton, S.** (2006). SUN1 interacts with nuclear lamin A
1012 and cytoplasmic nesprins to provide a physical connection between the nuclear
1013 lamina and the cytoskeleton. *Mol. Cell. Biol.* **26**, 3738–3751.

1014 **Hirsch, C. N., Hirsch, C. D., Brohammer, A. B., Bowman, M. J., Soifer, I., Barad, O.,**
1015 **Shem-Tov, D., Baruch, K., Lu, F., Hernandez, A. G., et al.** (2016). Draft Assembly
1016 of Elite Inbred Line PH207 Provides Insights into Genomic and Transcriptome
1017 Diversity in Maize. *Plant Cell* **28**, 2700–2714.

1018 **Hodzic, D. M., Yeater, D. B., Bengtsson, L., Otto, H. and Stahl, P. D.** (2004). Sun2 is
1019 a novel mammalian inner nuclear membrane protein. *J. Biol. Chem.* **279**, 25805–
1020 25812.

1021 **Horigome, C., Okada, T., Shimazu, K., Gasser, S. M. and Mizuta, K.** (2011).
1022 Ribosome biogenesis factors bind a nuclear envelope SUN domain protein to
1023 cluster yeast telomeres. *EMBO J.* **30**, 3799–3811.

1024 **Hotta, T., Haraguchi, T. and Mizuno, K.** (2007). A novel function of plant histone H1:
1025 microtubule nucleation and continuous plus end association. *Cell Struct. Funct.* **32**,
1026 79–87.

1027 **Jafferali, M. H., Vijayaraghavan, B., Figueroa, R. A., Crafoord, E., Gudise, S.,**
1028 **Larsson, V. J. and Hallberg, E.** (2014). MCLIP, an effective method to detect
1029 interactions of transmembrane proteins of the nuclear envelope in live cells.
1030 *Biochim. Biophys. Acta* **1838**, 2399–2403.

1031 **Janin, A., Bauer, D., Ratti, F., Millat, G. and Méjat, A.** (2017). Nuclear envelopathies:
1032 a complex LINC between nuclear envelope and pathology. *Orphanet J. Rare Dis.*
1033 **12**, 147.

1034 **Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., Campbell, M. S.,**
1035 **Stein, J. C., Wei, X., Chin, C.-S., et al.** (2017). Improved maize reference genome

1036 with single-molecule technologies. *Nature* **546**, 524–527.

1037 **Käll, L., Krogh, A. and Sonnhammer, E. L. L.** (2004). A combined transmembrane
1038 topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036.

1039 **Karimi, M., Inzé, D. and Depicker, A.** (2002). GATEWAY^{TM} vectors for Agrobacterium-
1040 mediated plant transformation. *Trends Plant Sci.* **7**, 193–195.

1041 **Kim, D. I., Birendra, K. C. and Roux, K. J.** (2015). Making the LINC: SUN and KASH
1042 protein interactions. *Biol. Chem.* **396**, 295–310.

1043 **Kopylov, M., Bass, H. W. and Stroupe, M. E.** (2015). The maize (Zea mays L.)
1044 nucleoside diphosphate kinase1 (ZmNDPK1) gene encodes a human NM23-H2
1045 homologue that binds and stabilizes G-quadruplex DNA. *Biochemistry* **54**, 1743–
1046 1757.

1047 **Kracklauer, M. P., Link, J. and Alsheimer, M.** (2013). LINCing the nuclear envelope to
1048 gametogenesis. *Curr. Top. Dev. Biol.* **102**, 127–157.

1049 **Krogh, A., Larsson, B., Von Heijne, G. and Sonnhammer, E. L. L.** (2001). Predicting
1050 transmembrane protein topology with a hidden markov model: application to
1051 complete genomes1. *J. Mol. Biol.* **305**, 567–580.

1052 **Kumar, S., Stecher, G. and Tamura, K.** (2016). MEGA7: Molecular Evolutionary
1053 Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874.

1054 **Li, P. and Noegel, A. A.** (2015). Inner nuclear envelope protein SUN1 plays a
1055 prominent role in mammalian mRNA export. *Nucleic Acids Res.* **43**, 9874–9888.

1056 **Lin, F., Jiang, L., Liu, Y., Lv, Y., Dai, H. and Zhao, H.** (2014). Genome-wide
1057 identification of housekeeping genes in maize. *Plant Mol. Biol.* **86**, 543–554.

1058 **Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., Lahrmann, U., Zhao, Q., Zheng, Y.,**
1059 **Zhao, Y., et al.** (2015). IBS: an illustrator for the presentation and visualization of
1060 biological sequences. *Bioinformatics* **31**, 3359–3361.

1061 **Luxton, G. W. G. and Starr, D. A.** (2014). KASHing up with the nucleus: novel

1062 functional roles of KASH proteins at the cytoplasmic surface of the nucleus. *Curr.*
1063 *Opin. Cell Biol.* **28**, 69–75.

1064 **Malone, C. J., Fixsen, W. D., Horvitz, H. R. and Han, M.** (1999). UNC-84 localizes to
1065 the nuclear envelope and is required for nuclear migration and anchoring during C.
1066 elegans development. *Development* **126**, 3171–3181.

1067 **Martinière, A., Lavagi, I., Nageswaran, G., Rolfe, D. J., Maneta-Peyret, L., Luu, D.-**
1068 **T., Botchway, S. W., Webb, S. E. D., Mongrand, S., Maurel, C., et al.** (2012). Cell
1069 wall constrains lateral diffusion of plant plasma-membrane proteins. *Proc. Natl.*
1070 *Acad. Sci. U. S. A.* **109**, 12805–12810.

1071 **Masuda, K., Xu, Z. J., Takahashi, S., Ito, A., Ono, M., Nomura, K. and Inoue, M.**
1072 (1997). Peripheral framework of carrot cell nucleus contains a novel protein
1073 predicted to exhibit a long alpha-helical domain. *Exp. Cell Res.* **232**, 173–181.

1074 **Meier, I.** (2016). LINCing the eukaryotic tree of life - towards a broad evolutionary
1075 comparison of nucleocytoplasmic bridging complexes. *J. Cell Sci.* **129**, 3523–3531.

1076 **Meier, I., Zhou, X., Brkljacić, J., Rose, A., Zhao, Q. and Xu, X. M.** (2010). Targeting
1077 proteins to the plant nuclear envelope. *Biochem. Soc. Trans.* **38**, 733–740.

1078 **Meier, I., Richards, E. J. and Evans, D. E.** (2017). Cell Biology of the Plant Nucleus.
1079 *Annu. Rev. Plant Biol.* **68**, 139–172.

1080 **Mochizuki, R., Tsugama, D., Yamazaki, M., Fujino, K. and Masuda, K.** (2017).
1081 Identification of candidates for interacting partners of the tail domain of DcNMCP1, a
1082 major component of the Daucus carota nuclear lamina-like structure. *Nucleus* **8**,
1083 312–322.

1084 **Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y. and Kurata, N.** (2005).
1085 Functional isolation of novel nuclear proteins showing a variety of subnuclear
1086 localizations. *Plant Cell* **17**, 389–403.

1087 **Murphy, S. P., Simmons, C. R. and Bass, H. W.** (2010). Structure and expression of
1088 the maize (Zea mays L.) SUN-domain protein gene family: evidence for the

1089 existence of two divergent classes of SUN proteins in plants. *BMC Plant Biol.* **10**,
1090 269.

1091 **Murphy, S. P., Gumber, H. K., Mao, Y. and Bass, H. W.** (2014). A dynamic meiotic
1092 SUN belt includes the zygotene-stage telomere bouquet and is disrupted in
1093 chromosome segregation mutants of maize (Zea mays L.). *Front. Plant Sci.* **5**, 314.

1094 **Nakayama, T., Ishii, T., Hotta, T. and Mizuno, K.** (2008). Radial microtubule
1095 organization by histone H1 on nuclei of cultured tobacco BY-2 cells. *J. Biol. Chem.*
1096 **283**, 16632–16640.

1097 **Nannas, N. J. and Dawe, R. K.** (2015). Genetic and genomic toolbox of Zea mays.
1098 *Genetics* **199**, 655–669.

1099 **Oda, Y. and Fukuda, H.** (2011). Dynamics of Arabidopsis SUN proteins during mitosis
1100 and their involvement in nuclear shaping. *Plant J.* **66**, 629–641.

1101 **Pawar, V., Poulet, A., Détourné, G., Tatout, C., Vanrobays, E., Evans, D. E. and**
1102 **Graumann, K.** (2016). A novel family of plant nuclear envelope-associated proteins.
1103 *J. Exp. Bot.* **67**, 5699–5710.

1104 **Poulet, A., Probst, A. V., Graumann, K., Tatout, C. and Evans, D.** (2017a). Exploring
1105 the evolution of the proteins of the plant nuclear envelope. *Nucleus* **8**, 46–59.

1106 **Poulet, A., Duc, C., Voisin, M., Desset, S., Tutois, S., Vanrobays, E., Benoit, M.,**
1107 **Evans, D. E., Probst, A. V. and Tatout, C.** (2017b). The LINC complex contributes
1108 to heterochromatin organisation and transcriptional gene silencing in plants. *J. Cell*
1109 *Sci.* **130**, 590–601.

1110 **Razafsky, D. and Hodzic, D.** (2015). Nuclear envelope: positioning nuclei and
1111 organizing synapses. *Curr. Opin. Cell Biol.* **34**, 84–93.

1112 **Shimada, N., Inouye, K., Sawai, S. and Kawata, T.** (2010). SunB, a novel Sad1 and
1113 UNC-84 domain-containing protein required for development of Dictyostelium
1114 discoideum. *Dev. Growth Differ.* **52**, 577–590.

1115 **Sohaskey, M. L., Jiang, Y., Zhao, J. J., Mohr, A., Roemer, F. and Harland, R. M.**
1116 (2010). Osteopotentia regulates osteoblast maturation, bone formation, and skeletal
1117 integrity in mice. *J. Cell Biol.* **189**, 511–525.

1118 **Sosa, B. A., Rothballer, A., Kutay, U. and Schwartz, T. U.** (2012). LINC complexes
1119 form by binding of three KASH peptides to domain interfaces of trimeric SUN
1120 proteins. *Cell* **149**, 1035–1047.

1121 **Sparkes, I. A., Runions, J., Kearns, A. and Hawes, C.** (2006). Rapid, transient
1122 expression of fluorescent fusion proteins in tobacco plants and generation of stably
1123 transformed plants. *Nat. Protoc.* **1**, 2019–2025.

1124 **Spichal, M. and Fabre, E.** (2017). The Emerging Role of the Cytoskeleton in
1125 Chromosome Dynamics. *Front. Genet.* **8**, 60.

1126 **Starr, D. A. and Fridolfsson, H. N.** (2010). Interactions between nuclei and the
1127 cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges. *Annu. Rev.*
1128 *Cell Dev. Biol.* **26**, 421–444.

1129 **Starr, D. A. and Han, M.** (2002). Role of ANC-1 in tethering nuclei to the actin
1130 cytoskeleton. *Science* **298**, 406–409.

1131 **Stelpflug, S. C., Sekhon, R. S., Vaillancourt, B., Hirsch, C. N., Buell, C. R., de Leon,**
1132 **N. and Kaeppler, S. M.** (2016). An Expanded Maize Gene Expression Atlas based
1133 on RNA Sequencing and its Use to Explore Root Development. *Plant Genome* **9**,.

1134 **Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T. and Hara-Nishimura, I.** (2010).
1135 Identification and characterization of nuclear pore complex components in
1136 Arabidopsis thaliana. *Plant Cell* **22**, 4084–4097.

1137 **Tamura, K., Iwabuchi, K., Fukao, Y., Kondo, M., Okamoto, K., Ueda, H., Nishimura,**
1138 **M. and Hara-Nishimura, I.** (2013). Myosin XI-i links the nuclear membrane to the
1139 cytoskeleton to control nuclear movement and shape in Arabidopsis. *Curr. Biol.* **23**,
1140 1776–1781.

1141 **Tamura, K., Goto, C. and Hara-Nishimura, I.** (2015). Recent advances in

1142 understanding plant nuclear envelope proteins involved in nuclear morphology. *J.*
1143 *Exp. Bot.* **66**, 1641–1647.

1144 **Tapley, E. C. and Starr, D. A.** (2013). Connecting the nucleus to the cytoskeleton by
1145 SUN-KASH bridges across the nuclear envelope. *Curr. Opin. Cell Biol.* **25**, 57–62.

1146 **Tiang, C.-L., He, Y. and Pawlowski, W. P.** (2012). Chromosome organization and
1147 dynamics during interphase, mitosis, and meiosis in plants. *Plant Physiol.* **158**, 26–
1148 34.

1149 **Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y.,**
1150 **Coppens, F. and Vandepoele, K.** (2018). PLAZA 4.0: an integrative resource for
1151 functional, evolutionary and comparative plant genomics. *Nucleic Acids Res.* **46**,
1152 D1190–D1196.

1153 **Varas, J., Graumann, K., Osman, K., Pradillo, M., Evans, D. E., Santos, J. L. and**
1154 **Armstrong, S. J.** (2015). Absence of SUN1 and SUN2 proteins in Arabidopsis
1155 thaliana leads to a delay in meiotic progression and defects in synapsis and
1156 recombination. *Plant J.* **81**, 329–346.

1157 **Vera, D. L., Madzima, T. F., Labonne, J. D., Alam, M. P., Hoffman, G. G.,**
1158 **Girimurugan, S. B., Zhang, J., McGinnis, K. M., Dennis, J. H. and Bass, H. W.**
1159 (2014). Differential nuclease sensitivity profiling of chromatin reveals biochemical
1160 footprints coupled to gene expression and functional DNA elements in maize. *Plant*
1161 *Cell* **26**, 3883–3893.

1162 **Walley, J. W., Sartor, R. C., Shen, Z., Schmitz, R. J., Wu, K. J., Urich, M. A., Nery, J.**
1163 **R., Smith, L. G., Schnable, J. C., Ecker, J. R., et al.** (2016). Integration of omic
1164 networks in a developmental atlas of maize. *Science* **353**, 814–818.

1165 **Wang, H., Dittmer, T. A. and Richards, E. J.** (2013). Arabidopsis CROWDED NUCLEI
1166 (CRWN) proteins are required for nuclear size control and heterochromatin
1167 organization. *BMC Plant Biol.* **13**, 200.

1168 **Wang, B., Tseng, E., Regulski, M., Clark, T. A., Hon, T., Jiao, Y., Lu, Z., Olson, A.,**

1169 **Stein, J. C. and Ware, D.** (2016). Unveiling the complexity of the maize
1170 transcriptome by single-molecule long-read sequencing. *Nat. Commun.* **7**, 11708.

1171 **Xu, X. M., Meulia, T. and Meier, I.** (2007). Anchorage of plant RanGAP to the nuclear
1172 envelope involves novel nuclear-pore-associated proteins. *Curr. Biol.* **17**, 1157–
1173 1163.

1174 **Xu, X. H., Chen, H., Sang, Y. L., Wang, F., Ma, J. P., Gao, X.-Q. and Zhang, X. S.**
1175 (2012). Identification of genes specifically or preferentially expressed in maize silk
1176 reveals similarity and diversity in transcript abundance of different dry stigmas. *BMC*
1177 *Genomics* **13**, 294.

1178 **Zhao, Q., Brkljacic, J. and Meier, I.** (2008). Two distinct interacting classes of nuclear
1179 envelope-associated coiled-coil proteins are required for the tissue-specific nuclear
1180 envelope targeting of Arabidopsis RanGAP. *Plant Cell* **20**, 1639–1651.

1181 **Zhou, X., Graumann, K., Evans, D. E. and Meier, I.** (2012). Novel plant SUN–KASH
1182 bridges are involved in RanGAP anchoring and nuclear shape determination. *J. Cell*
1183 *Biol.* **196**, 203–211.

1184 **Zhou, X., Graumann, K., Wirthmueller, L., Jones, J. D. G. and Meier, I.** (2014).
1185 Identification of unique SUN-interacting nuclear envelope proteins with diverse
1186 functions in plants. *J. Cell Biol.* **205**, 677–692.

1187 **Zhou, X., Graumann, K. and Meier, I.** (2015a). The plant nuclear envelope as a
1188 multifunctional platform LINCed by SUN and KASH. *J. Exp. Bot.* **66**, 1649–1659.

1189 **Zhou, X., Groves, N. R. and Meier, I.** (2015b). Plant nuclear shape is independently
1190 determined by the SUN-WIP-WIT2-myosin XI-i complex and CRWN1. *Nucleus* **6**,
1191 144–153.

1192 **Zhou, X., Groves, N. R. and Meier, I.** (2015c). SUN anchors pollen WIP-WIT
1193 complexes at the vegetative nuclear envelope and is necessary for pollen tube
1194 targeting and fertility. *J. Exp. Bot.* **66**, 7299–7307.

1195

1196

Table 1: Maize LINC Genes

	Gene Identifier[A]					Protein Sequence Features					Domains		
Maize Gene name[B]	**UniProt[C]**	**B73v3[D]**	**B73v4[D]**	**PLAZA gene[E]**	**Plaza gene family[E]**	**N-terminal Sequence**	**C-terminal Sequence**	**Total length (AA)**	**Molecular weight (kDa)**	**pI[G]**	**CC**	**TMD**	**Other (location)**
MLKG1	A0A096TG70	GRMZM2G370707	Zm00001d002723	ZM02G25900	HOM04M009187	MDMPP	VVPT*	258	28	4.82	74- 94	224 - 246	KASH (247 - 258)
MLKG2	A0A1D6M739	GRMZM2G117392	Zm00001d038539	ZM06G27180	HOM04M008346	MTKPT	VMPPT*	227	23.74	5.02	-	191 - 214	KASH (215 - 227)
MLKS1	A0A1D6KGN6	GRMZM2G094850	Zm00001d031134	ZM01G30030	HOM04M002474	MATSL	VVPT*	617	67.2	8.59	-	579 - 601	KASH (602-617)
MLKS2	K7UCH8	GRMZM2G000608	Zm00001d052955	ZM04G32760	HOM04M002474	MGRSL	LVPT*	637	68.49	6.55	-	605 - 623	KASH (624 - 637), ARM (3 - 517)
MLKP1	B4FDR2	GRMZM2G151418	Zm00001d003334	ZM02G12590	HOM04M002181	MSMNY	AVPT*	294	33.73	4.6	68 - 91, 109 -136, 149 - 183, 222 - 256	264 - 286	KASH (289 - 294)
MLKP2	K7TL67	GRMZM2G034882	Zm00001d025667	ZM10G18600	HOM04M002181	MLMHY	VVPT*	290	33.57	4.63	68 - 91, 109 -136, 152 -179	258 - 280	KASH (281 - 290)
MLKP3	B7ZXA8	GRMZM2G003992	Zm00001d005997	ZM02G32490	HOM04M002181	MDSGA	IVPT*	538	57.9	5.47	319 - 343, 358 - 378, 452 - 500	506 - 529	KASH (530 - 538)
MLKP4	C0PH11	GRMZM2G176407	Zm00001d020879	ZM07G17460	HOM04M002181	MDSGA	IVPT*	532	57.6	4.7	313 - 337, 352 - 372, 401 - 421, 446 - 480	502 - 523	KASH (524 - 532)
MLKT1	COP4P9	GRMZM2G148723	Zm00001d039272	ZM03G00380	HOM04M000144	MDFHN	VNDC*	653	72.35	4.86	94-114, 135-155, 201-228, 374-415, 455-517	620-640	KASH (641 - 653)
MLKT2	K7UYZ4	GRMZM2G006661	Zm00001d010047	ZM08G14110	HOM04M000144	MDTVI	GSPA*	667	74.51	4.72	96-116, 130-157, 218-238, 316-392, 411-431, 453-473, 481-508, 527-547	643-661	KASH (662 - 667)
ZmSUN1	B6THU8	GRMZM2G109818	Zm00001d015450	ZM05G20890	HOM04M002694	MTMSA	PSTGG*	462	50.96	9.07	169 - 189, 195 - 229	116 - 141	SUN (277-455)
ZmSUN2	B6TY16	GRMZM2G440614	Zm00001d040331	ZM03G08780	HOM04M002694	MASPS	AGSQA*	439	47.5	7.78	168-188	84 - 109	SUN (264-478)
ZmSUN3	A0A1D6N5R4	GRMZM2G120262	Zm00001d042643	ZM03G25290	HOM04M001731	MQRSR	ILIIQ*	613	68.04	4.86	484 - 511	33 - 55, 554 - 574, 595 -612	SUN (194 - 358)
ZmSUN4	K7VAL9	GRMZM2G005483	Zm00001d012240	ZM08G30590	HOM04M001731	MSLSC	ILIIQ*	639	70.54	5.22	510 -537	57 - 79, 580 -600, 621 -638	SUN (218 - 382)
ZmSUN5	A0A1D6FYJ8	AC194341.4_FG003	Zm00001d011277	ZM08G23230	HOM04M001731	MSRKR	VLLYN*	607	65.86	5.46	414 - 434, 495 - 523	46 - 66, 525 -544, 572 - 588	SUN (152 - 322)
MNEAP1	B6TAT2	GRMZM2G042593	Zm00001d010918	ZM08G20470	HOM04M002962	MSVSE	VILAKR*	348	39.75	8.94	221 - 248	330-346	NA
MNEAP2	K7TRE8	GRMZM2G061728	Zm00001d024756	ZM10G11720	HOM04M002962	MQCLS	VILAKR*	347	39.73	8.97	149 - 83 220 – 247	329-345	NA
MNEAP3	B6TCY4	(absent)	Zm00001d002875	(absent)	HOM04M002962	MSLLK	FIHV*	170	19.89	8.94	14 – 34, 84-104	125-150	NA
NCH1	K7U7Q6	GRMZM2G015875	Zm00001d051600	ZM04G22450	HOM04M004640	MFTPQ	SFFTT*	1156	132.82	5.13	160 – 278, 312 – 350, 362 – 522, 537 – 579, 592 – 701	NA	
NCH2	A0A1D6NAM9	GRMZM2G320013	Zm00001d043335	ZM03G30810	HOM04M003564	MASPR	KMFKE*	896	105.04	5.25	131 – 165, 250 – 281, 309 – 336, 386 – 532, 635 – 721	NA	
MKAKU41	A0A1D6JGB5	GRMZM2G104199	Zm00001d026487	ZM10G25400	HOM04M004991	MASLF	RGRAT*	554	59.59	5.29	NA	NA	
MKAKU42	A0A1D6DVM2	GRMZM2G521862	Zm00001d002012	ZM02G01940	HOM04M004991	MASLF	EGRAT*	591	64.87	5.66	NA	NA	

FOOTNOTES

A Gene IDs are listed as of 2017

B New gene names MKLP, MLKS, Maize LINC KASH AtSINE-like; MLKG, Maize LINC KASH Grass-specific; Maize LINC KASH AtWIP-like: MLKT, Maize LINC KASH AtWIT-like

C UniProt ID from the UnitProt Consortium as of December 2017 (Nucleic Acids Res. 45: D158-D169 (2017))

D B73v3 and B73v4 gene models from MaizeGDB.org as of December 2017

E Plaza gene and family IDs from version 4 of Plaza Gene (https://doi.org/10.1093/nar/gkx1002)

F KASH domain refers to amino acids post transmembrane domain

G pI, Isoelectric point calculated from the full-length protein using ExPASy Compute "pI/Mw"

Table 2: Proteins co-precipitated with ZmSUN2 pull down on maize ear shoot nuclei

	B73v4 Maize gene id	Maize gene name	Gene Description[1]	Exclusive unique peptide count	SUN IP (Normalized total spectra)	Irrelevant antibody control IP	No Antibody control IP
LINC components	Zm00001d040331	ZmSUN2	SUN domain protein2	11	384.45	0	0
	Zm00001d015450	ZmSUN1	SUN domain protein1	3	78.398	0	0
	Zm00001d051600	NCH1	Protein CROWDED NUCLEI 1	2	82.803	0	0
	Zm00001d043335	NCH2	Nuclear lamina component	3	28.871	0	0
	Zm00001d002723	MLKG1	OSJNBa0010H02.21-like protein	2	14.459	0	0
	Zm00001d052955	MLKS2	ARM repeat superfamily protein	2	1.2886	0	0
	Zm00001d010047	MLKT2	WPP domain-interacting tail-anchored protein 1	1	26.117	0	0
Nucleoporins	Zm00001d052282	NUA	Nuclear-pore anchor	7	87.161	0	0
	Zm00001d048023	NUP133	Nuclear pore complex protein NUP133	1	18.911	0	0
	Zm00001d046566	NUP96	Nuclear pore complex protein NUP96	1	7.253	0	0
	Zm00001d026263	NUP88	Nuclear pore complex protein NUP88	1	7.206	0	0
	Zm00001d005355	GP210	Nuclear pore complex protein GP210	1	7.206	0	0
Cytoskeletal elements	Zm00001d013410	Actin-1	ARP7 Nuclear actin	3	82.756	0	0
	Zm00001d012556	Tub6	beta- Tubulin 6	5	40.576	0	0
	Zm00001d015348	Tub4	beta- Tubulin 4	1	7.253	0	0
Ribosomal proteins	Zm00001d044287	L19	Ribosomal protein L19	2	40.623	0	0
	Zm00001d017047	L6	60S ribosomal protein L6	1	30.569	0	0
	Zm00001d022463	L22-2	60S ribosomal protein L22-2	1	11.658	0	0
RAN	Zm00001d017239	RanGAP2	RAN GTPase-activating protein 2	1	14.459	0	0
	Zm00001d048223	RABA6a	Ras-related protein RABA6a	2	14.412	0	0
DNA/RNA binding	Zm00001d052820	RNA binding protein	RNA-binding (RRM/RBD/RNP motifs) family protein	2	21.665	0	0
	Zm00001d025638	bZIP5	Basic-leucine zipper domain	1	11.658	0	0
	Zm00001d042267	ARF10	Auxin response factor 10	1	11.658	0	0
	Zm00001d011297	MYBR35	Putative MYB DNA-binding domain superfamily protein	1	7.206	0	0
	Zm00001d016354	bHLH72/bHLH041	Putative transcription factor bHLH041	1	7.206	0	0
Histone	Zm00001d034479	H1a	Histone 1 (H1)	1	26.117	0	0

Figure 1

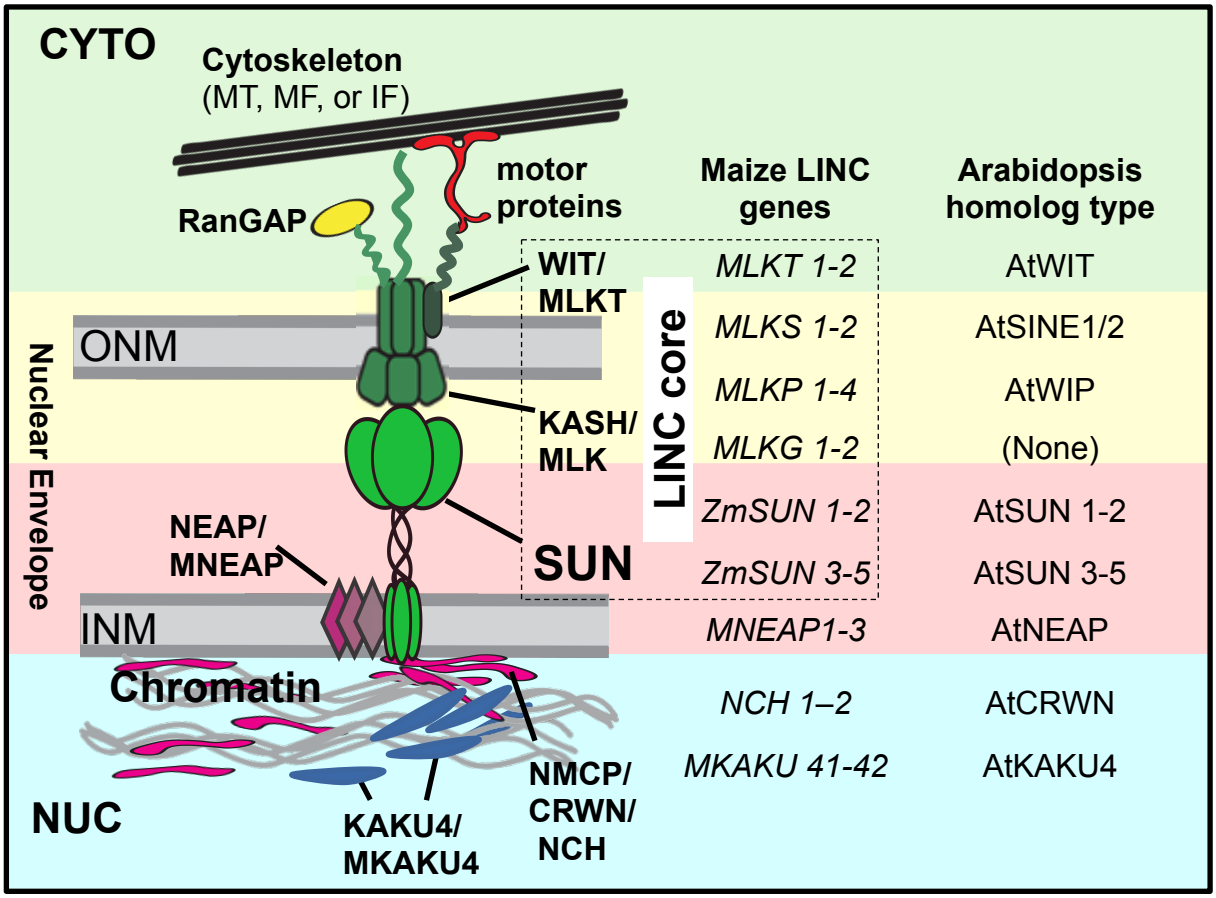

CYTO
Cytoskeleton
(MT, MF, or IF)
motor
proteins
RanGAP
Maize LINC
genes
Arabidopsis
homolog type
WIT/
MLKT
MLKT 1-2
AtWIT
ONM
MLKS 1-2
AtSINE1/2
MLKP 1-4
AtWIP
KASH/
MLK
MLKG 1-2
(None)
LINC core
Nuclear Envelope
ZmSUN 1-2
AtSUN 1-2
NEAP/
MNEAP
SUN
ZmSUN 3-5
AtSUN 3-5
INM
MNEAP1-3
AtNEAP
Chromatin
NCH 1–2
AtCRWN
MKAKU 41-42
AtKAKU4
NMCP/
CRWN/
NCH
NUC
KAKU4/
MKAKU4

Figure 2

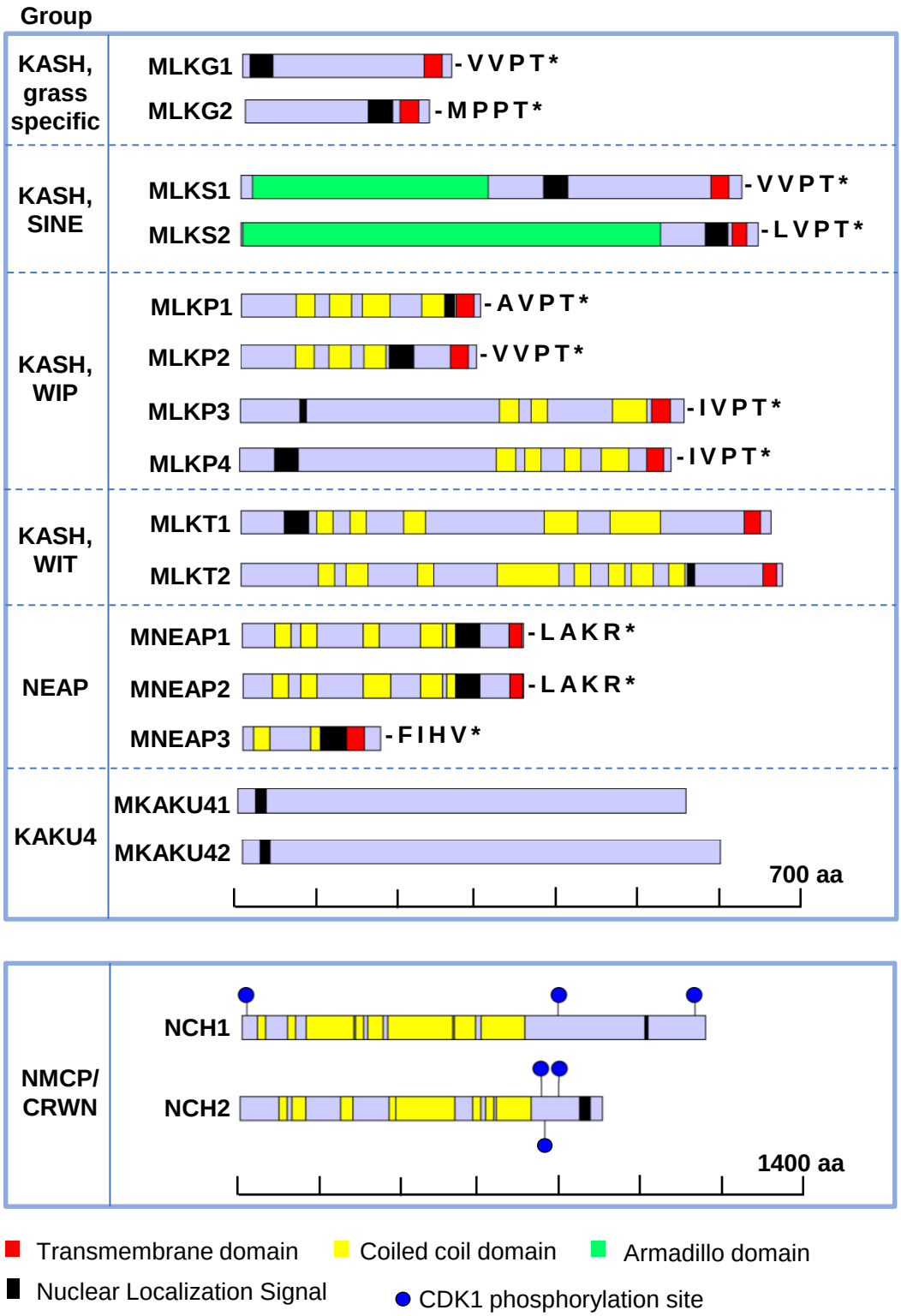

Group
KASH, grass specific
MLKG1
-VVPT*
MLKG2
-MPPT*
KASH, SINE
MLKS1
-VVPT*
MLKS2
-LVPT*
KASH, WIP
MLKP1
-AVPT*
MLKP2
-VVPT*
MLKP3
-IVPT*
MLKP4
-IVPT*
KASH, WIT
MLKT1
MLKT2
NEAP
MNEAP1
-LAKR*
MNEAP2
-LAKR*
MNEAP3
-FIHV*
KAKU4
MKAKU41
MKAKU42
700 aa
NMCP/ CRWN
NCH1
NCH2
1400 aa
Transmembrane domain
Coiled coil domain
Armadillo domain
Nuclear Localization Signal
CDK1 phosphorylation site

Figure 3

General grouping color key

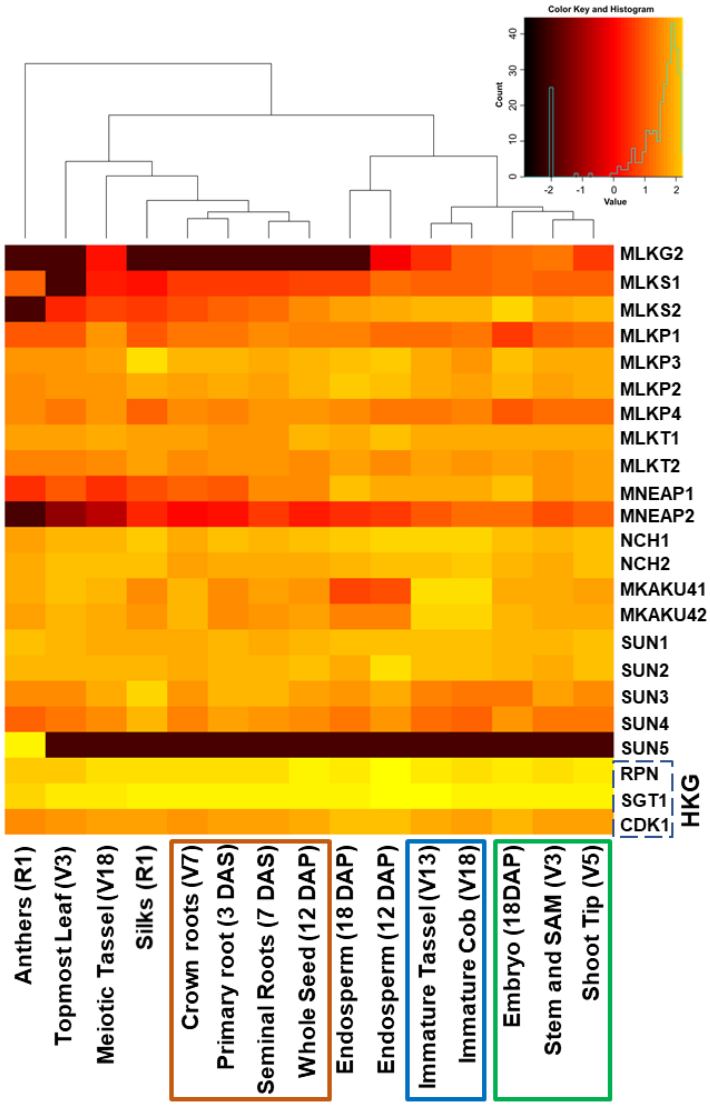
- Green plants
- Land plants
- Seed plants
- Flowering plants
- Grass family

Viridiplantae (Green plants)
Land plants
Seed plants
Angiosperms (Flowering plants)
Monocots
Eudicots
Poaceae (Grasses)
Gymno
Bryophyte
Algae

Gene names	Location	Group	Plaza gene family (HOMO4M)	Zema	Sobi	Seit	Hovu	Trae	Brdi	Orsa	Zoma	Arth	Potr	Vivi	Soly	Metr	Piab	Phpa	Chre
MLKG1	ONM	KASH, grass	009187	1	1	1	1	3	1	1	0	0	0	0	0	0	0	0	0
MLKG2	ONM	KASH, grass	008346	1	1	1	1	6	1	0	0	0	0	0	0	0	0	0	0
MLKS1- MLKS2	ONM	KASH, SINE	002474	2	2	2	1	6	2	2	1	2	2	1	1	1	2	7	0
MLKP1 – MLKP4	ONM	KASH, WIP	002181	4	3	3	2	7	3	3	1	3	2	1	1	2	0	0	0
MLKT1-MLKT2	ONM	WIT	000144	2	1	2	2	5	1	2	2	2	1	2	2	1	0	0	0
ZmSUN1- ZmSUN2	INM	C-term SUN	002694	2	2	2	2	6	2	2	2	2	2	1	1	1	1	2	0
ZmSUN3- ZmSUN5	INM	mid-SUN	001731	3	2	2	2	6	2	2	3	3	3	2	1	3	2	2	1
MNEAP1 – MNEAP3	INM	NEAP	002962	3	2	2	1	2	1	2	2	4	2	2	1	1	1	0	0
NCH1	NUC	NMCP/CRWN	004640	1	1	1	2	3	1	1	4	1	2	2	2	2	2	0	0
NCH2	NUC	NMCP/CRWN	003564	1	1	1	1	3	1	1	4	3	3	1	1	1	2	2	0
MKAKU41- MKAKU42	NUC	KAKU4	004991	2	1	1	1	3	1	1	2	1	2	1	1	1	0	0	0

Figure 4

A


Color Key and Histogram
Count
Value
MLKG2
MLKS1
MLKS2
MLKP1
MLKP3
MLKP2
MLKP4
MLKT1
MLKT2
MNEAP1
MNEAP2
NCH1
NCH2
MKAKU41
MKAKU42
SUN1
SUN2
SUN3
SUN4
SUN5
RPN
SGT1
CDK1
HKG
Anthers (R1)
Topmost Leaf (V3)
Meiotic Tassel (V18)
Silks (R1)
Crown roots (V7)
Primary root (3 DAS)
Seminal Roots (7 DAS)
Whole Seed (12 DAP)
Endosperm (18 DAP)
Endosperm (12 DAP)
Immature Tassel (V13)
Immature Cob (V18)
Embryo (18DAP)
Stem and SAM (V3)
Shoot Tip (V5)

B

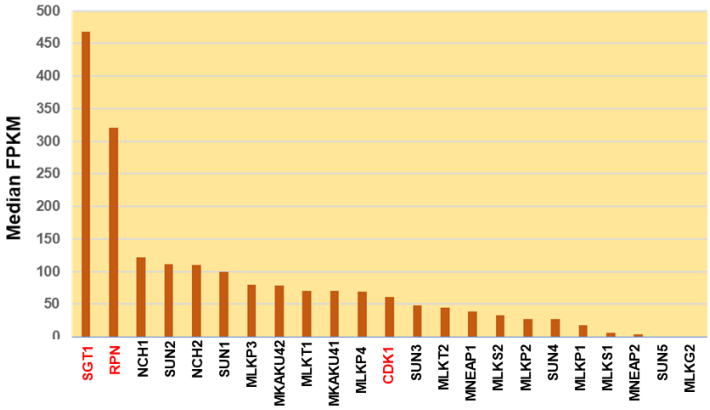

Median FPKM
500
450
400
350
300
250
200
150
100
50
0
SGT1
RPN
NCH1
SUN2
NCH2
SUN1
MLKP3
MKAKU42
MLKT1
MKAKU41
MLKP4
CDK1
SUN3
MLKT2
MNEAP1
MLKS2
MLKP2
SUN4
MLKP1
MLKS1
MNEAP2
SUN5
MLKG2

Figure 5

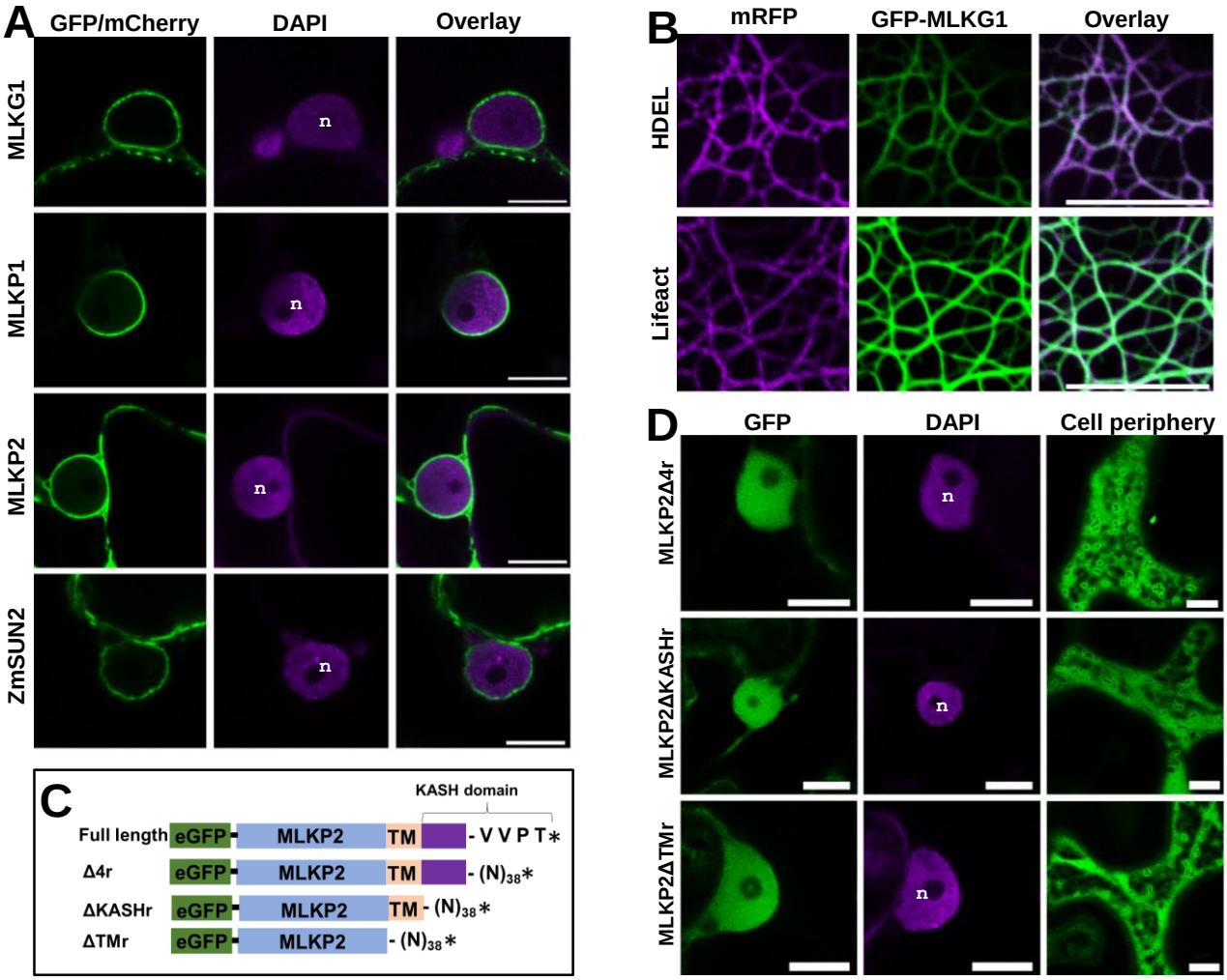
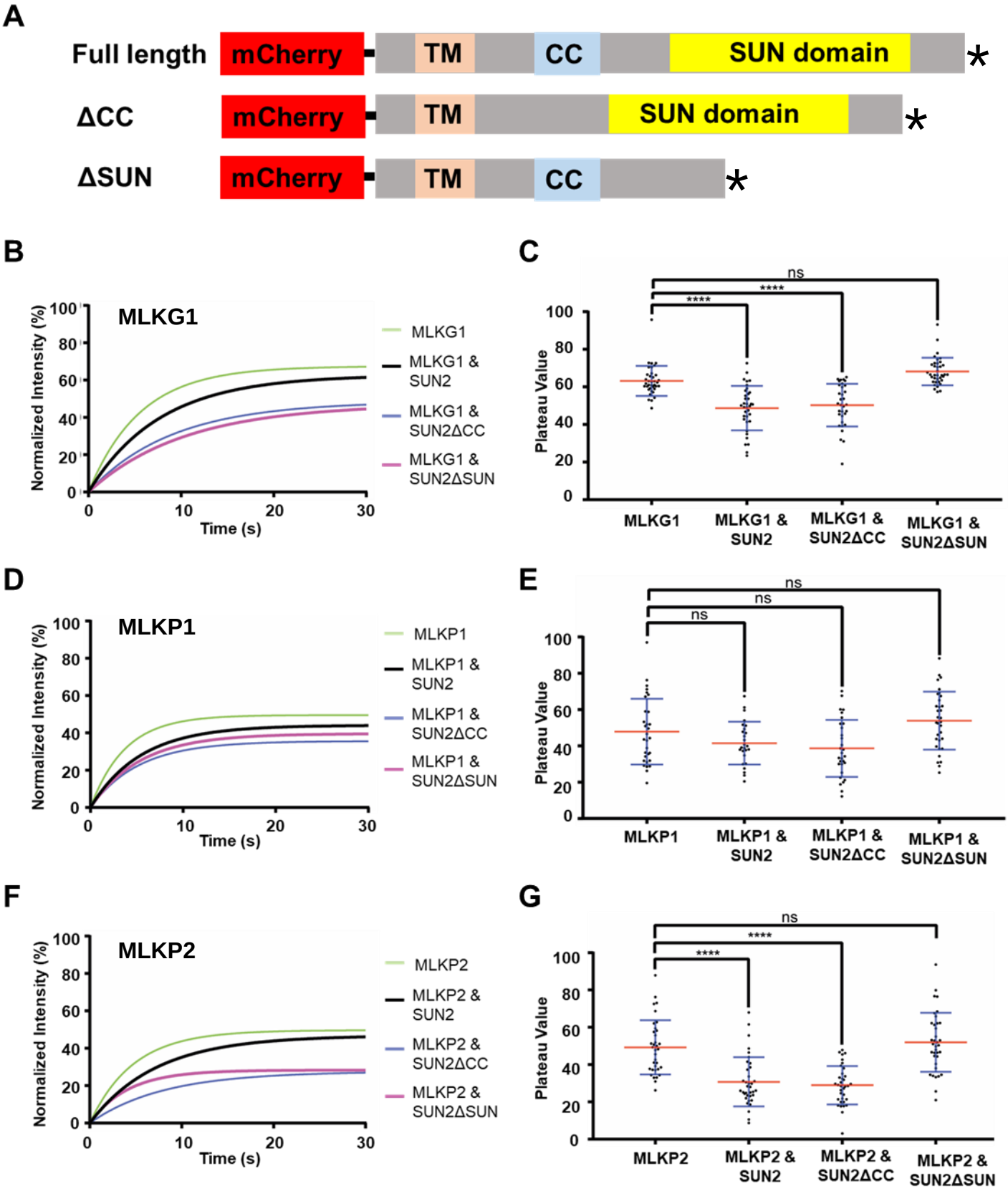

A
GFP/mCherry
DAPI
Overlay
MLKG1
MLKP1
MLKP2
ZmSUN2
n
B
mRFP
GFP-MLKG1
Overlay
HDEL
Lifeact
C
KASH domain
Full length
eGFP
MLKP2
TM
- V V P T*
Δ4r
- (N)38*
ΔKASHr
ΔTMr
D
GFP
DAPI
Cell periphery
MLKP2Δ4r
MLKP2ΔKASHr
MLKP2ΔTMr

Figure 6


A
Full length
mCherry
TM
CC
SUN domain
*
ΔCC
mCherry
TM
SUN domain
*
ΔSUN
mCherry
TM
CC
*
B
MLKG1
Normalized Intensity (%)
Time (s)
MLKG1
MLKG1 & SUN2
MLKG1 & SUN2ΔCC
MLKG1 & SUN2ΔSUN
C
Plateau Value
ns
****
****
MLKG1
MLKG1 & SUN2
MLKG1 & SUN2ΔCC
MLKG1 & SUN2ΔSUN
D
MLKP1
Normalized Intensity (%)
Time (s)
MLKP1
MLKP1 & SUN2
MLKP1 & SUN2ΔCC
MLKP1 & SUN2ΔSUN
E
Plateau Value
ns
ns
ns
MLKP1
MLKP1 & SUN2
MLKP1 & SUN2ΔCC
MLKP1 & SUN2ΔSUN
F
MLKP2
Normalized Intensity (%)
Time (s)
MLKP2
MLKP2 & SUN2
MLKP2 & SUN2ΔCC
MLKP2 & SUN2ΔSUN
G
Plateau Value
ns
****
****
MLKP2
MLKP2 & SUN2
MLKP2 & SUN2ΔCC
MLKP2 & SUN2ΔSUN