

#### **Supplemental Figure 1**

Large scale and higher magnification view of localization of KNL2(C), KNL2(C)<sup>CENPC-k(R-A)</sup> and KNL2(C)<sup>CENPC</sup> fused to EYFP in root tips of stable transformants of *A. thaliana*. Corresponding small scale view is presented on Figure 2.



## Supplemental Figure 2

The CENPC motif of CENP-C protein restores the ability of KNL2(C)<sup> $\Delta CENPC$ </sup> to localize at centromeres.

Localization of native KNL2(C) (A), mutagenized KNL2 with deleted CENPC-k motif (KNL2(C)<sup> $\Delta CENPC-k$ </sup>) (B), and mutagenized KNL2 with CENPC-k motif substituted by the CENPC motif of the CENP-C protein (KNL2(C)<sup>CENPC</sup>) (C) fused to EYFP in nuclei of *Nicotiana* benthamiana transiently expressing the corresponding constructs. To visualize centromeres, all three constructs were infiltrated together with a construct expressing a cenH3-mCherry fusion protein.



### **Supplemental Figure 3**

# The C-terminal part of KNL2 binds the centromeric pAL1 sequences also in radioactive EMSA.

Five different radioactive EMSA experiments with 0,2 pmol labeled *pAL1* (all lanes) without (-) or with (+) 1-2 pmol purified recombinant C-terminal KNL2. Complexes of *pAL1* and KNL2(C) remain close to the slots similar to non-radioactive EMSA experiments.

А



#### **Supplemental Figure 4**

#### KNL2 DNA binding is not influenced by centromeric 23 nt ssRNA.

(A) EMSA experiment with 1 pmol of 23 nt ssRNA (all lanes) and reduced amounts of purified recombinant C-terminal KNL2 (1; 0.8; 0.6; 0.4; 0.2; 0.1; 0.05 pmol in the lanes 2-9, respectively).

(B) Three different radioactive EMSA experiments with 0.1 pmol labeled *pAL1* (all lanes) and 1 pmol purified recombinant C-terminal KNL2 (lanes 2-6). Increasing amounts of 23 nt ssRNA of *pAL1* were added (lane 3 - 0.1 pmol, lane 4 - 1 pmol. lane 5 - 10 pmol, lane 6 - 100 pmol). ssRNA does not interfere the KNL2(C)-pAL1 complex formation and stability.

Sequence: Prediction:	MTEPNLDEDGSKSSFQKTVVLRDWWLIKCPKEFEGKQFGVAGFEESVETRAMRVFTSSPI
Confidence:	5147797655200021026793856892762456725569997771752145 <mark>9</mark> 5410015
Sequence: Prediction: Confidence:	TKALDVFTLLASDGIYITL <mark>R</mark> GFLNKERVLKNGFNPEISREFIFGFPPCWERVCNSCFEGD
	1268798397837672706 <mark>9</mark> 6485150561468332720569787876350665368688
Sequence: Prediction: Confidence:	SFGTDVNTVPSTIEKACPPILSPCKYSN <mark>R</mark> NLKDNPAES <mark>R</mark> EKSNVTETDIAEINDKGGSGA
	3772563183215309554651220000 <mark>9</mark> 161432640 <mark>9</mark> 211251428999957145134
Sequence: Prediction: Confidence:	RDIKTA <mark>RRR</mark> SLHLQIKRILESSKV <mark>R</mark> KTANDGDHGSEFLNTAKRGDVERDGCEVINNEDSE + <mark>+++</mark>
	<mark>9</mark> 23004 <mark>999</mark> 042547115630105 <mark>9</mark> 01446765835573262063782889899648936
Sequence: Prediction: Confidence:	WKLDESEVQNLCNDGDNGSEGFIKAKSSDVEKDKSEAIDNDVISPAVGSGIKHTGADNVD
	639875695698777746235571501227528235899799924775145022577677
Sequence: Prediction: Confidence: Sequence: Prediction: Confidence:	KVTSASATGESLTSEQQNGLLVTTASPHSLLKDLAKSS <mark>K</mark> PEKKGIS <mark>K</mark> KSGKILRSDDNVV
	17014040310411435467550050321661657100 <mark>9</mark> 0300240 <mark>9</mark> 0030561386699
	DPMNYSGTKVKSAENKRKIDASKLQSPTSNVAEHSKEGLNNAKSNDVEKDVCVAINNEVI
	86541020 <mark>9</mark> 4002222005251 <mark>9</mark> 4002012554411257326012389599999956689
Sequence: Prediction: Confidence:	SPVKGFGKRLSGTDVERLTSKNATKESLTSVQRKGRVKVSKAFQDPLSKGKSKKSEKTLQ
	249053600514257307100030 <mark>9</mark> 1051051003 <mark>9</mark> 304105635361 <mark>9298999</mark> 19031
Sequence: Prediction:	SNSNVVEPMNHFRSEAEEAEENLSWEKIKRKIDFDVEVTPEKKVKQQKTNAASTDSLGQK
Confidence:	141378776536045899987682220711278999791330030010112301115301
Prediction:	RSRSGRVLVSSLEFWRNQIPVYDMDRNLIQVKDGSETNSAPSKGRGSDSRKRRNLKIK
conflaence:	atan2a221111200143808318331134010422010228a1a1a00a3aa02040
*** Prediction:	binding residues are labeled with '+' and in red; non-binding residues labeled with '-' and in green.
*** Confidence:	from level 0 (lowest) to level 9 (highest).

#### Supplemental Figure 5

*In silico* analysis of KNL2 protein by BindN program (Wang and Brown 2006) for DNA binding amino acid residues

To check for a possible DNA binding region in KNL2 the BindN program was used. Multiple residues with high DNA binding probability (red frame) arise at the C-terminus around the CENPC-k motif (blue frame).

# Supplemental Table 1

# Primers used in this study.

Primer name	Primer sequence	Reference
Side-directed mutagenesis	-	
KNL2 R A f	GTCAAGATCAGGAGCGGTGCTTGTGTCATC	
KNL2 R A r	GTTTCTGTCCCAATGAATCAGTAGACGCCGC	
KNL2_W_R_f	CATCACTAGAGTTTCGGCGTAACCAAATTC	
KNL2_W_R_r	ACACAAGCACCCTTCCTGATCTTGACCGTT	
KNL2_ACENPC_f	GATGGTAGTGAGACTAACTCCGCTC	
KNL2_ACENPC_r	TCCCAATGAATCAGTAGACGCCGC	
CENP-Csubst_fr1	GCATCCATGAGAGTTTGACTACTGTTGATGGT	
	AGTGAGACTAACTCCGCTC	
CENP-Csubst_rev1	GTCCGTACAAGAATCTTTCACCTCTTCCCAAT	
	GAATCAGTAGACGCCGC	
CENP-Csubst_fr2	TTAAGTCGAGACCGCTCGAATACTGGAGAGGT	
	GAAAGATTCTTGTACGGAC	
CENP-Csubst_rev2	TCCTTGTACTTCGTCTGACACCACCTCCCAAT	
	GAATCAGTAGACGCCGC	
Generation of DNA samples		
Atnila_I	ATCACGTTGCCGACATCATAACG	
Atnila_r		(Maria tana 1, 2000)
Telomere_I		(Yu et al., 2000)
Telomere_r		(Yu et al., 2000)
	GCTGTTGGTGGAGGGACTGG	(Ay et al., 2009)
		(May et al., 2009)
PALI_I		(May et al., 2005)
PALI_I	IIGCIICICAAAGAIIICAIGGI	(May et al., 2005)
ChIP and gPCR		
Athila f (0.1uM.)*	ATCACGTTGCCGACATCATAACG	
Athila r (0.1uM)	AGCAACAACGAAAGAGTTGAGAGC	
Subtelo f (0.25µM)	CTAAACTAGTTGTGTTCCCGTCTCTACT	(Gallego and White, 2001)
Subtelo r (0.25µM)	GGTGGGCGACCTTGTGCTTGCCAAAGTC	(Gallego and White, 2001)
TUA4 f (0.04µM)	GCTGTTGGTGGAGGGACTGG	(Ay et al., 2009)
TUA4 r (0.04µM)	CCTGTGGAGATGGGTAAACTGTG	(Ay et al., 2009)
pAL1 f (0.05µM)	GGTTAGTGTTTTGGAGTCGAATATG	(May et al., 2005)
pAL1 r (0.05µM)	TTGCTTCTCAAAGATTTCATGGT	(May et al., 2005)
serva		
23 bp ssRNA of pAL1	ACAAGGAUACAAUUCUUACGCCU	
forward pall seRNA	CCULIAUCUUUUCCACUCCAAUAUCACUUCAUC	(7hand et al 2008)
IOIWAIG PAIL SSIGNA		(2mang et al., 2000)
	IIGAGAAGCAA	
reverse pAll SSRNA		(7hangetal 2008)
TOTOLOG PHILI SOUNA	GUCCAUAUGAGUCUUGGCUUUGUGUGUGUGUGUGUGUGUG	(inding cc ar., 2000)
	GUUGCGUUUUUAAGUUCUUAUACUCAAUCAUAC	
	ACAUGAGAUCGAUCAAGUCAUAUUUCGACUCCA	
	AAACAUAACC	

\*-final concentration of primers

#### Supplemental Table 2

DNA	PCR program
fragment	
Athila	94°C – 5 min / (94°C – 20 s, 54°C – 30 s, 72°C – 1 min)x45 / 72°C – 7 min /
	4°C - ∞
TUA4	94°C – 5 min / (94°C – 20 s, 60°C – 30 s, 72°C – 1 min)x35 / 72°C – 7 min /
	4°C - ∞
Telomere	94°C – 5 min / (94°C – 20 s, 60°C – 30 s, 72°C – 5 s)x35 / 4°C - ∞
pAL1	94°C – 5 min / (94°C – 20 s, 50°C – 30 s, 72°C – 3 s)x40 / 4°C - ∞

#### PCR programs used for the amplification of DNA probes for EMSA.

#### REFERENCES

- Ay, N., Irmler, K., Fischer, A., Uhlemann, R., Reuter, G., and Humbeck, K. (2009). Epigenetic programming via histone methylation at WRKY53 controls leaf senescence in Arabidopsis thaliana. Plant J 58, 333-346.
- Gallego, M.E., and White, C.I. (2001). RAD50 function is essential for telomere maintenance in Arabidopsis. P Natl Acad Sci USA **98**, 1711-1716.
- May, B.P., Lippman, Z.B., Fang, Y., Spector, D.L., and Martienssen, R.A. (2005). Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. Plos Genet 1, e79.
- Yu, E.Y., Kim, S.E., Kim, J.H., Ko, J.H., Cho, M.H., and Chung, I.K. (2000). Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. J Biol Chem 275, 24208-24214.
- **Zhang, W., Lee, H.R., Koo, D.H., and Jiang, J.** (2008). Epigenetic modification of centromeric chromatin: hypomethylation of DNA sequences in the CENH3-associated chromatin in Arabidopsis thaliana and maize. The Plant cell **20**, 25-34.