Structural requirements for CENH3 targeting to centromeric chromatin

Dissertation

Submitted to the Faculty of Mathematics, Natural Sciences and Technology at the Martin-Luther University Halle-Wittenberg in partial fulfilment of the requirements for the degree of

doctor rerum naturalium (Dr. rer. nat.)

by

Izabel C. R. Moraes

born on February 25, 1979, in Sao Paulo, Brazil

Reviewers:

Prof. Dr. Gunter Reuter

Prof. Dr. Stephan Diekmann

Prof. Dr. Ingo Schubert

Halle (Saale), February/2011

ACKNOWLEDGEMENT

I would like to thank Prof. Ingo Schubert for giving me the opportunity to be part of the karyotype evolution group and to develop my PhD work here, for the continuous support and for the very helpful suggestions especially in writing this thesis.

Many thanks to Inna Lermontova for the discussions, scientific advices and for keeping a nice atmosphere in our office. Spasiba!

Thanks to the members of KTE group, especially Stefan Heckmann and Jörg Fuchs, for the nice discussions about chromosomes, seasons, german cakes, music, and other important subjects.

My gratitude to Joachim, Rita, Andrea and Martina for technical support and to Inge for helping with all the administrative issues.

I'm very thankful to Twan Rutten and Diaa Dadghma for helping me with use of the confocal microscope.

Many thanks to Gabriele Brantin (Transcriptome Analysis) for assistance with the transient expression studies.

I also thank Prof. Gunter Reuter, Prof. Stephan Diekmann and Prof. Ingo Schubert for accepting to make part of the committee for my PhD work evaluation.

Special thanks to my friends outside the lab for the nice moments even under snow and/or -20ºC!

Os meus mais sinceros agradecimentos à minha família: minha mãe, Marina, e meu pai, Francisco, minhas irmãs Denise e Sandra, meu "irmão" Sandro e meu sobrinho Lucas, pelo apoio, incentivo e encorajamento em todos os momentos.

Obrigada ao Nico Blanco, por trazer luz a dias nublados.

À Dra. Vera Quecini, pelas discussões e orientações que tanto me ajudaram no curso desses três anos, o meu muitíssimo obrigada!

Content

5. DISCUSSION

List of figures

List of tables

List of abbreviations

35S-pro	Cauliflower mosaic virus promoter
A. tumefaciens	*Agrobacterium tumefaciens*
Amp	Ampicilin
BAP	6-Benzylaminopurine
BB	Blocking buffer
bp	Base pairs
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediamine tetraacetic acid
EYFP	Enhanced Yellow Fluorescent Protein
GFP	Green Fluorescent Protein
GM	Germination medium
kb	Kilo Base Pairs
kDa	Kilo Dalton
LB	Lysogeny broth
M	Molarity
Min	Minute(s)
ml	Millilitre
mM	Millimol
MTSB	Microtubules Stabilizing Buffer
µ	Micro
MS	Murashige & Skoog medium
pmol	Picomol
PPT	Phosphinotrycin
PCR	Polymerase chain reaction
PEG	Polyethylenglycol
RNA	Ribonucleic acid
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
Sec	Second(s)

SDS	Sodiumdodecylsulfat
SOC	Super Optimal Broth With Catabolite Repression
SSC	Sodium Chloride Sodium Citrate Buffer
T1	First generation of plants after transformation
T2	Second generation of plants after transformation
Taq	*Thermus aquaticus*
Tris	2-Amino-2-hydroxymethyl-1,3-propandiol
WT	Wild type
YEB	Yeast Extract Broth

1. Introduction

1.1 Eukaryotic chromosome organization

Mitosis is the essential process by which eukaryotic cells divide, resulting in two daughter cells carrying the same genetic component. The genetic information in all the forms of life and some viruses is carried by a nucleic acid molecule, the deoxyribonucleic acid (DNA).

In eukaryotes, the DNA is divided in a set of chromosomes. Within the chromosome, the long linear DNA molecule is folded and packed by proteins into a compact structure. This complex of DNA and associated proteins is called chromatin (Alberts et al. 2002).

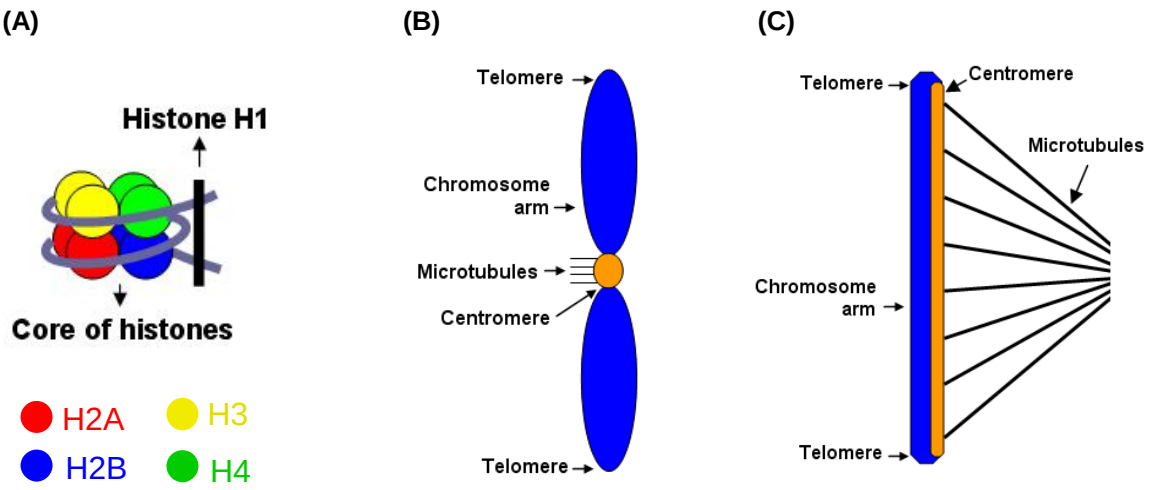

FIGURE 1 Chromosome organization A) the nucleosome structure is the basic chromatin unit, composed of core histones and DNA (purple) linked by histone H1; domains of a B) typical eukaryotic monocentric chromosome, and C) of a holocentric chromosome with the kinetochore distributed throughout its entire length.

Nucleosomes represent the basic unit of chromatin. Within the core of a nucleosome, the DNA is wrapped around an octamer of proteins called histones. Histones are small basic proteins containing a central core region, the histone fold domain, and C- and N-terminal tails. Each nucleosome contains two copies of histones H2A, H2B, H3 and H4 (Fig. 1 A). The linker histone H1 connects individual nucleosomes forming the 30 nm chromatin fiber which is organized in loops (reviewed in Heng et al. 2001) attached to a scaffold of non-histone proteins. During metaphase the chromatin folds into its maximally compacted structure.

According to their particular features, chromatin has been classified into euchromatin and heterochromatin. Euchromatin is marked by unique coding sequences and transcriptional activity (reviewed in Gill et al. 2008). In contrast, heterochromatin is highly condensed, contains few genes, and is enriched with highly repeated (satellite) DNA. Despite its transcriptional inactivity, heterochromatin plays an important roles in chromosome inheritance, genome stability, and dosage compensation in animals (reviewed in Lam et al. 2005).

In order to promote or prevent transcription, chromatin structure can be modified. These modifications occur either on heterochromatin or euchromatin and include post-translational modification of canonical histones, substitution of histone variants and remodeling of chromatin (reviewed in Francis 2009). Post-translational modifications of canonical histones usually occur at the N-terminal tails, and include acetylation, phosphorylation, ubiquitination, and methylation (reviewed in Lam et al. 2005).

Heterochromatin is commonly found in three chromosome domains: centromeres, telomeres and nuclear organizing regions (NORs) (McCombie et al. 2000). Most eukaryotic centromeres and telomeres are composed of tandem arrays of repetitive sequences, while NORs consist mostly of ribosomal RNA genes (5.8S, 18S and 28S) (reviewed in Gill et al. 2008). Telomeres are the terminal region of linear eukaryotic chromosomes specified by conserved DNA sequences and proteins (Lamb et al. 2007). While telomeres guarantee terminal stability to the chromosome, centromeres are required for their faithful segregation.

1.2 Centromeres

A chromosome component that has an important role during mitosis and meiosis is the centromere (Fig. 1 B-C). The centromere, consisting of centromeric DNA and a protein complex, the kinetochore, is constitutively present during the cell cycle and responsible for chromatid cohesion, spindle fiber attachment and chromosome movement (reviewed in Houben and Schubert 2003).

Chromosomes are classified according to centromere localization into monocentric and holocentric. Monocentric chromosomes have one size-restricted centromere per chromosome which appears as the primary constriction. In contrast, in holocentric chromosomes the centromere is dispersed over nearly the entire length of the chromosome (Fig. 1 C), therefore no primary constriction is visible. In monocentric chromosomes the microtubules attach to a distinct kinetochore, moving the sister chromatids to the pole during anaphase with the centromere leading. In holocentrics, nearly the entire length of the chromosome is attached by the microtubules, moving the chromatids as a linear bar to the pole (Nagaki et al. 2005). The best studied holocentric chromosomes among plants are those of the monocot genus *Luzula.* Among Metazoan some insects, arachnids and nematodes have holocentric chromosomes, of which the best studied are those of the nematode *Caenorhabditis elegans*.

Centromeres vary in size and sequence composition (Fig. 2), from the very small 125 bp point centromeres of budding yeast, consisting of specific DNA sequences (Clarke 1990), to the several megabases of regional centromeres of humans, composed of repetitive DNA (Morris and Moazed 2007).

The DNA sequence within the point centromere of *S. cerevisae* has the sub-domains CDE-I, CDE-II and CDE-III (Centromere DNA Element I, II and III). Deletions within CDEI (Hegemann et al. 1988) and CDEII (Sears et al. 1995) affect chromosome segregation and single point mutations within CDEIII abolish centromere function (Jehn et al. 1991, McGrew et al. 1986).

In contrast to point centromeres, which are composed of specific DNA sequences important for protein binding, regional centromeres are much larger

and their activity does not depend on a specific DNA sequence (Choo 2001). Indeed, the variability of centromere sequences is very high (Fig. 2).

Schizosaccharomyces pombe centromeres contain central cores of DNA (non-repetitive central core [cnt] and centromere specific inter repeats [imr]) surrounded by long tandem repeats (ort) (Baum et al. 1994). While otr has a conserved sequence in all the three chromosomes of *S. pombe*, cnt and imr are variable (Ishii 2009). Human centromeres are composed of ~171-bp repeats, the α-satellites (Schueler et al. 2001). In *Drosophila*, centromeres are composed of small repeats of only 5 bp, interspersed with transposable elements (Sun et al. 1997).

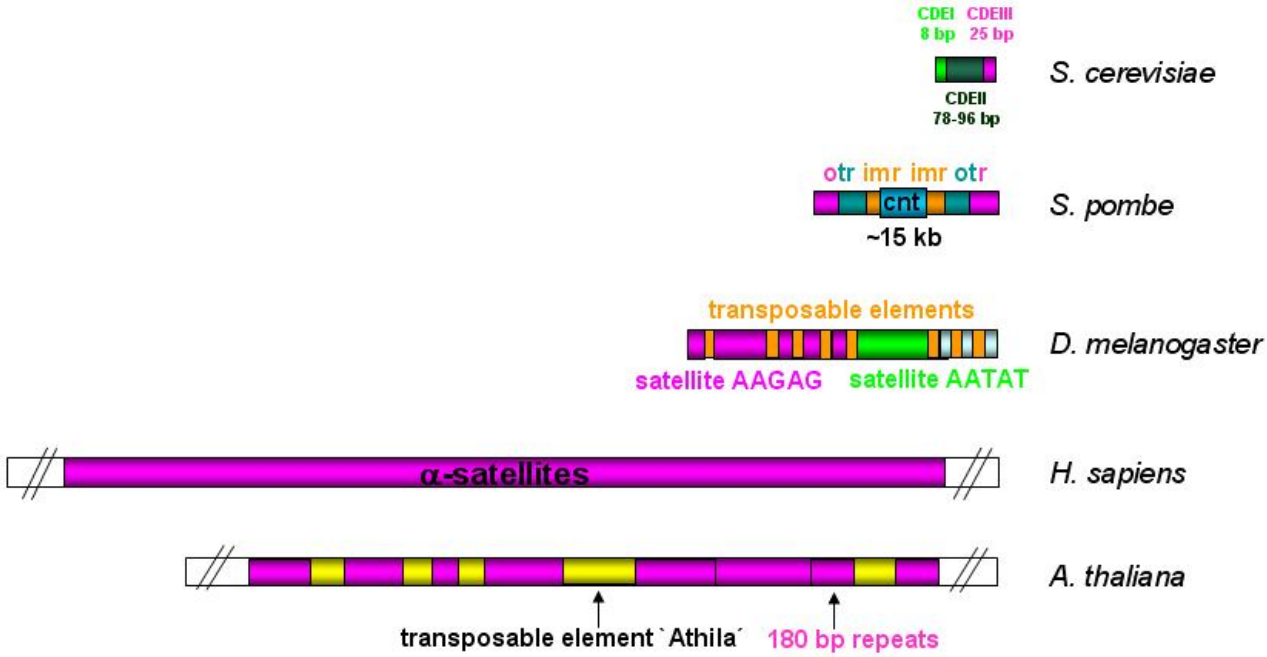

FIGURE 2 High variability of centromeric DNA. From top to bottom: the simple point centromere of the budding yeast *S. cerevisiae* (~125 bp), regional centromeres of fission yeast *S. pombe* (~40-100 kb), *Drosophila melanogaster* (~420 kb), *H. sapiens* (250-5000 kb), and the thale cress *A. thaliana* (~4000 kb)

Most plants have regional centromeres. In *A. thaliana,* the centromeric region comprises ~180-bp satellite repeats (Martinez-Zapater et al. 1986) and the

internal portions of Athila2 (May et al. 2005), Athila LTR-retroelements (Hall et al. 2003) and 106B repeats (Thompson et al. 1996).

Centromeric DNA composition has been elucidated also in some cereals (reviewed in Houben and Schubert 2003). Two conserved centromeric sequences, cereal centromeric sequence (CCS1) (Aragón-Alcaide 1996) and Sau3A9 (Jiang et al. 1996) are present in wheat, rye, barley, maize and rice. These repetitive sequences were found to be parts of the LTR (Long Terminal Repeat) and of the integrase region of the Ty*3*/*gypsy*-like retroelement *cereba* (centromeric retrotransposon of barley) respectively (Presting et al. 1998).

1.3 Kinetochore structure and assembly

Kinetochore is a complex structure of more than 90 proteins which assembles at the centromeric position of each chromosome during mitosis. A combination of epigenetic marks, chromatin structure and DNA sequence properties determine the centromere as the site for kinetochore assembly (reviewed in Gascoigne and Cheeseman 2010).

Kinetochore structure considerably varies between organisms (Fig. 3). The vertebrate kinetochore has a trilaminar structure composed of an outer layer and an inner layer on the surface of the centromeric heterochromatin (Fig. 3 A-B). Observations on electron microscopy revealed that the two domains are separated by a translucent middle zone (Chan et al. 2005). In the absence of microtubules, a complex of fibers called fibrous corona can be visualized on the surface of the outer plate (Ris and Witt 1981). This structure is assembled in mitosis, when interaction with microtubules is necessary, and is absent during interphase (Gascoigne and Cheeseman 2010). In contrast, plant kinetochores (Fig. 3 C-D) are not layered and appear as an inclusion set into the surface of the chromosome, so-called “ball in the cup” configuration (reviewed in Baskin and Cande 1990). The name ‘ball in the cup’ refers to the uniform kinetochore ultrastructure embedded in a “cup” of chromatin (Yu et al. 2000) distinguishable from the cytoplasm by the attached set of microtubules (Dawe et al. 2005).

In contrast to the variable centromeric DNA sequences, kinetochore proteins are conserved (reviewed in Cooper and Henikoff 2004; Malik and Henikoff 2001). For instance, kinetochore proteins such as ZW10, CENH3, CENP-C and

BUB1 were identified in *C. elegans* due to homology to those found in monocentric chromosomes (reviewed in Dernburg 2001).

Kinetochore proteins have been classified into constitutive proteins, associated with centromeres throughout the cell cycle, and transient proteins, commonly found on outer kinetochore for interaction with spindle microtubules (Cheeseman and Desai 2008).

The first kinetochore proteins identified were CENP-A, CENP-B and CENP-C (reviewed in Cheeseman and Desai 2008). The antibodies able to recognize these proteins were isolated from patients that developed a variant of scleroderma, CREST (calcinosis, Raynauds's phenomenon, esophageal dismotility, sclerodactyly, telangiectasia) (Moroi et al. 1980). Sera from CREST patients also recognize plant kinetochore proteins, providing a strong evidence of their high conservation (Yu et al. 2000). Two of these proteins are present in all eukaryotes: CENP-A (Centromere Protein A, Earnshaw and Rothfield 1985, called generally CENH3, Centromeric Histone H3) and CENP-C, both localize in the inner kinetochore and are only found in active centromeres, but not on inactive ones (Choo 1997). Inhibition of CENP-C by either antibody microinjection or RNAi-mediated depletion resulted in failure of kinetochore assembly (Tomkiel et al. 1994; Kwon et al. 2007). In Drosophila, CENP-C is also required for DmCENH3 targeting to the centromere (Erhardt et al. 2008). Three CENP-C homologues were found constitutively present at *Z. mays* centromeres and co-localizing with CENH3 in meiotic cells (Dawe et al. 1999; Zhong et al. 2002).

Although Ogura (2004) observed AtCENP-C constitutively present in dividing cultured cells of *A. thaliana*, recent findings revealed that CENP-C is not detectable in differentiated nuclei (Lermontova et al. submitted). In accordance to this data, during development of *Drosophila* the amount of CENP-C decreases after mitosis, while in cells proliferating mitotically the signals were intense (Heeger et al. 2005).

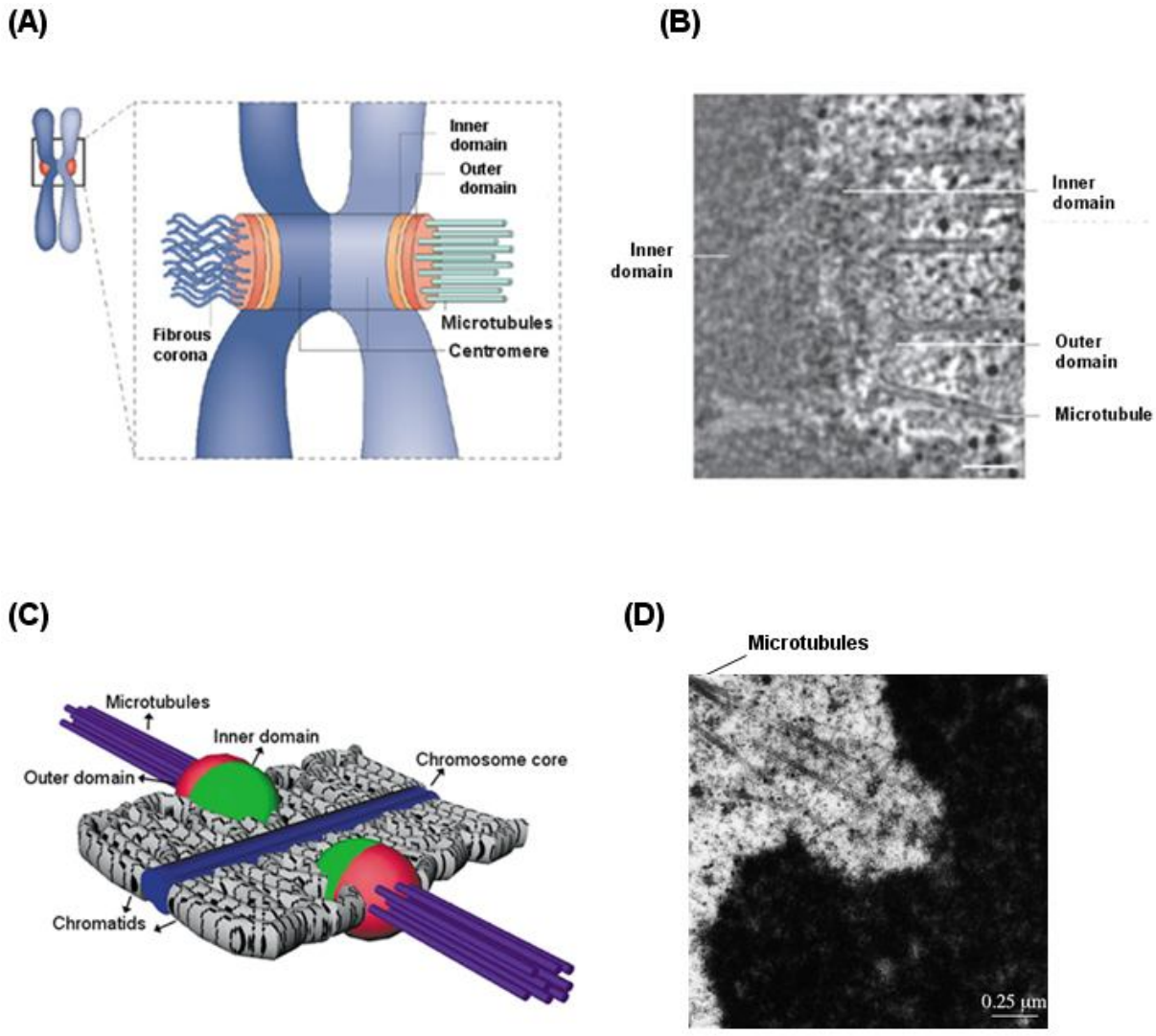

FIGURE 3 Kinetochore types and domains **(A)** Scheme of the trilaminar kinetochore, commonly found in chromosomes of vertebrate (from Cheeseman and Desai 2008) composed of inner and outer domains. In the absence of microtubules, a complex of fibers called fibrous corona can be visualized on the surface of the outer plate; **(B)** electron micrograph of a human kinetochore (from Cheeseman and Desai 2008); bar = 100 nm **(C)** scheme of the plant kinetochore, referred to as a 'ball in the cup' (from Yu et al. 2000) and **(D)** ultrastructure of *Z. mays* kinetochores, here in prometaphase II (from Dawe et al. 2005); the "ball" refers to the kinetochore embedded in a "cup" of chromatin.

Transient proteins of the outer kinetochore include motor, chromatin modifying and checkpoint proteins. Motor proteins are involved in chromosome movement and spindle checkpoint control. They include the kinesin superfamily and CENP-E protein, which plays an essential role in capturing and positioning chromosomes at metaphase plate (Schaar et al. 1997). Two putative homologues of CENP-E were identified in plants: cpeI1 and cpeI2 of the monocot *Hordeum vulgare* and the dicot *Vicia faba* (ten Hoopen et al. 2002).

The *S*pindle *A*ssembly *C*heckpoint (SAC) is the mechanism that ensures faithful chromosome segregation during mitosis. The SAC regulates the activation of the ubiquitin ligase complex known as anaphase promoting complex (APC), preventing transition to anaphase until all the chromosomes achieve bipolar attachment. The inhibitory “wait anaphase” signal is generated from unattached kinetochores or lack of spindle tension, and leads to bounding of the co-factor CDC20 by MAD and/or BUB complexes, preventing the activation of APC (Sudakin et al. 2001). When all sister kinetochores are attached in a bipolar orientation to spindle fibers, the tension causes release of checkpoint proteins and dissociation of their complexes with CDC20. As an activator of APC, CDC20 promotes ubiquitination and degradation of securin, what results in activation of separase for separation and poleward segregation of sister chromatids (Robbins and Cross 2010).

The SAC components BUB (budding uninhibited by benzimidazole) and MAD (mitotic arrest-deficient) were identified from a genetic screen in yeast to isolate mutants insensitive to spindle damage (Li and Murray 1991; Hoyt et al. 1991; Murray 1992). While the SAC is in metazoans essential to prevent chromosome mis-segregation and cell death (Musacchio and Salmon 2007), in budding yeast is non-essential, being important only in response to perturbation of the attachment process (Li and Murray 1991, Hoyt et al. 1991). In plants, the SAC proteins include MAD2 of wheat (Kimbara et al. 2004), maize (Yu et al. 1999) and *A. thaliana* (Caillaud et al. 2009), BUB3 (Lermontova et al. 2008), and BUBR1 (Caillaud et al. 2009) of *A. thaliana*.

Although kinetochore can bind to microtubules in any configuration, correct chromosome segregation requires that each kinetochore attaches to microtubules from opposite spindle poles (reviewed in Lampson and

Cheeseman 2010). Errors during this process, which could result in unequal segregation, are avoided by a specialized protein that stabilizes the attachment through tension across the centromere (Nicklas and Koch 1969; Nicklas and Ward 1994). This conserved protein was identified in a screen of budding yeast mutants showing an increase-in-ploidy (Ipl1, Francisco et al. 1994). Ipl1 was shown to phosphorylate kinetochore substrates, regulating microtubule binding (Biggins et al. 1999). Aurora kinases, homologues of Ipl1, were identified in *D. melanogaster* (Glover et al. 1995)*, C. elegans (*Schumacher et al. 1998), H. sapiens (Bischoff et al. 1998) and *A. thaliana* (Demidov et al. 2005).

To form attachment sites for the spindle microtubules, three conserved proteins associate to the kinetochore: 1) KNL-1 (kinetochore null), 2) the Mis12 complex and 3) the Ndc80 complex (Santaguida and Musacchio 2009). Homologues in plants include an *A. thaliana* Mis12 (Sato et al. 2005) and NDC80 of maize (Du and Dawe 2007), both constitutively associated to the chromatin throughout the cell cycle. KNL-1 proteins were not detected so far in plants.

The constitutive centromere-associated network (CCAN), a group of proteins including CENP-C, -H, -I, -K, -U and -W, is present in the centromere throughout the cell cycle and required for kinetochore assembly. The CCAN directs KNL-1 protein and Mis12 complex, which recruits the Ndc80 complex during kinetochore assembly (reviewed in Cheeseman and Desai 2008)

Although kinetochores vary in organization, composition and size, all of them assemble on centromeric nucleosomes containing a modified H3, the Centromeric histone H3 (CENH3). CENH3 replaces histone H3 in centromeric nucleosomes and is required for kinetochore formation.

1.3.1 The Centromeric Histone H3

CENH3 consists of a highly variable N-terminal part and a more conserved C-terminal part containing the loop1 region (Fig. 4), conserved among *Arabidopsis* species, but less so between more distantly related species (Malik and Henikoff 2002).

Since it is present in centromeric nucleosomes CENH3 serves as an epigenetic mark for the centromeres. In metazoans, CENH3 is incorporated during

anaphase/telophase to G1. In yeast it occurs during S and G2, and in plants and protozoans during G2 (Table 1).

Budding yeast CENH3 containing nucleosomes were previously reported as octameric (Camahort et al. 2009) in contrast to the tetrameric centromeric nucleosomes found in *Drosophila* (Dalal et al. 2007). Recently, Dimitriadis et al. (2010) detected tetrameric CENH3 containing nucleosomes *in vivo* in human cells. In centromeric chromatin of *Drosophila, C. elegans* and mammalians, CENH3 containing nucleosomes are interspersed with H3 containing nucleosomes, originating a structure in which CENH3 is positioned to the poleward of chromosome. This structure is though to facilitate the recruitment of additional kinetochore proteins and bi-orient sister kinetochores (Blower et al. 2002).

The loop1 region of the C-terminal part is essential for centromere targeting of CENH3 in *A. thaliana* (Lermontova et al. 2006). Cells of plants carrying EYFP-tagged N-terminal part had dispersed signals over the entire nuclei, while EYFP fused to C-terminal part was accumulated at chromocenters.

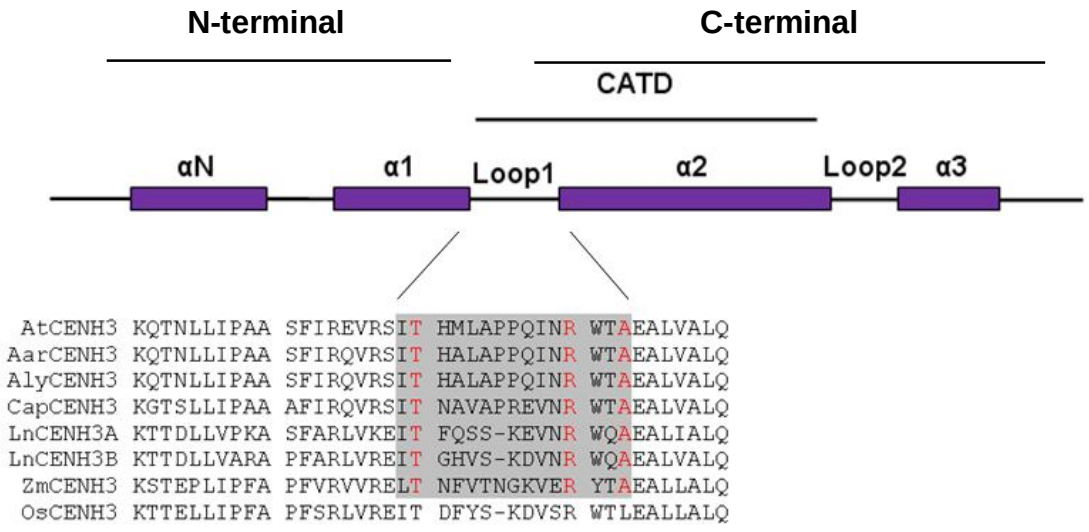

FIGURE 4 Centromeric Histone H3 domains and sequence alignment of the loop1 region of different plant species. The first 4 species are dicots, and the last 3 are monocots. *Luzula nivea* (Ln) is a species with holocentric chromosomes.

The region defined as CENP-A targeting domain (CATD) within the histone fold domain (HFD) of the C-terminal part is composed of the a2-helix and loop1 (Fig. 4). Recombinant human histone H3 carrying CENH3 CATD is also targeted to the centromeres (Black et al. 2004, 2007).

Table 1 CENH3 homologues, deposition time and chaperones

Organism	Homologue	Deposition during cell cycle	Chaperone / assembly complex	References
H. sapiens	CENP-A	Telophase / early G1	RbAp48 HJURP	Earnshaw and Rothfield 1985 Palmer et al. 1991 Jansen et al. 2007 Hemmerich et al. 2008 Foltz et al. 2009 Dunleavy et al. 2009
Drosophila	Cid	Anaphase, telophase	RbAp48	Ahmad and Henikoff 2001 Schuh et al. 2007
S. cerevisiae	Cse4	S phase	Mis16 (RbAp48) Mis18 Scm3	Meluh et al. 1998 Stoler et al. 2007
C. elegans	HCP-3 CPAR-I		KLN-2 (Mis18BP)	Buchwitz et al. 1999 Monen et al. 2007 Maddox et al. 2004 Moore et al. 1999
S. pombe	Cnp1	S/G2	Mis16 Mis18 Sim3 Scm3	Hayashi et al. 2004 Takahashi et al. 2005
A. thaliana	AtCENH3	Late G2		Talbert et al. 2004 Lermontova et al. 2006
Hordeum vulgare	HvCENH3	Late G2		Lermontova et al. 2007
Luzula nivea	LnCENH3	Prophase to Metaphase		Nagaki et al. 2005
Cyanidioschyzon merolae	CmCENH3	S to metaphase		Maruyama et al. 2007
Dyctiostellium	DdCENH3	G2		Dubin et al. 2010

In *Drosophila* the conservation of the loop1 region is critical for centromere targeting. DmCENH3 carrying the loop 1 sequence of *D. bipectinata* could not target the *D. melanogaster* centromeres, but *D. bipectinata* CENH3 with the loop1 region of DmCENH3 could (Veermak et al. 2002).

The N-terminal part, dispensable for targeting, seems to play a role for protein-protein interactions (Vermaak et al. 2002, Chen et al. 2000). In budding yeast centromeres, the interaction between the protein complex Ctf19p-Mcm21p-Okp1p with CENH3 is abolished in mutants with a deletion at CENH3 N-terminal part (Chen et al. 2000).

Mice heterozygous for CENH3 null mutation are healthy and fertile, while in homozygous state is embryo lethal (Howman et al. 2000). *Drosophila* embryos homozygous for null mutation show delay in early mitotic stages: the mutants had an increase in the number of prophase and prometaphase, and very few cells were able to progress to anaphase. Although the SAC response is not affected, some kinetochore proteins including SAC components do not target to kinetochores (Blower et al. 2006).

CENH3 is required for recruitment of many centromere proteins for kinetochore assembly, for instance those composing the CCAN (Cheeseman and Desai 2008). Consequently, many kinetochore components are not recruited during mitosis in cells lacking CENH3 (Amor et al. 2004).

Despite its conserved function CENH3 evolves rapidly in *Arabidopsis* and *Drosophila* (Malik and Henikoff 2001, Talbert et al. 2002, Cooper and Henikoff 2004), particularly its N-terminal domain (Malik and Henikoff 2001).

1.3.2 CENH3 recruitment and loading factors

Nucleosome assembly is a two-steps reaction, starting with 1) the deposition of a tetramer of histones H3 and H4 onto DNA followed by 2) two H2A–H2B heterodimers, resulting in an octamer core. Histone chaperones bind and neutralize the highly positive charge of the histones (Philpott et al. 2000, Furuyama et al. 2006).

The best characterized chaperone is the chromatin assembly factor 1 (CAF-1), which mediates the first step of nucleosome assembly (Smith and Stillman 1989, 1991; Shibahara and Stillman 1999; Tagami et al. 2004). Human CAF-1

contains the subunits p150, p60 and p48 (Ridgway and Almouzni 2000). *A. thaliana* CAF-1 subunits are encoded by the genes Fas1 (FASCIATA1), Fas2 (FASCIATA2), and Msi1 (multicopy suppressor of *ira1),* respectively (Kaya et al. 2001; Hennig et al. 2003).

Another important chaperone is the anti-silencing factor 1 (Asf1) that forms a complex with H3 and H4 and assists chromatin assembly and remodeling during replication, transcription activation, and gene silencing (English et al. 2006).

Proteins guiding CENH3 targeting have been identified and characterized in yeast, *Drosophila*, mammals and *C. elegans* (Table 1) (Silva and Jansen 2009, Dalal and Bui 2010). While canonical H3 is deposited during replication and is the substrate for CAF-1, CENH3 incorporates to chromatin independently of DNA replication (Shelby et al. 2000; Ahmad and Henikoff 2001). Apparently CENH3 is loaded through different pathways than the canonical H3.

In fission yeast, Mis16 and Mis18 are required for proper loading of Cnp1 (Fujita et al. 2007, Hayashi et al. 2004). hMis18 depletion resulted in CENP-A localization at non-centromeric regions, showing that also in humans Mis18 is required for CENP-A association with centromeres (Fujita et al. 2007). Similarly, in *C. elegans*, depletion of KLN-2 (mis18 homologue in worms) resulted in ectopic localization of CENH3 (Maddox et al. 2007). RbAp46 and RbAp48 (Mis15 and Mis16 homologues in human) are required for CENP-A targeting in yeast, suggesting that the chromatin-remodeling complex influences CENP-A deposition (Dunleavy et al. 2009).

In *S. cerevisiae*, a Suppressor of Chromosome Missegregation (Scm3) is required for CENH3 recruitment to the centromeres (Pidoux et al. 2009). Because the interaction of Scm3 with CENH3 depends on Sim3 (Pidoux et al. 2009), Sim3 has been proposed to be a chaperone that delivers CENH3 to the assembly factors (Dunleavy et al. 2007).

The recently identified HJURP (Holliday junction recognizing protein) (Foltz et al. 2009) was proposed to be a chaperone which deliveries, incorporates and maintains CENP-A at the centromeres. HJURP associates with CENP-A before its assembly and is recruited to the centromeres in vertebrates exactly at the

same time when newly synthesized CENP-A is deposited, i.e. during telophase to early G1 (Dunleavy et al. 2009).

Although loading time of CENH3 and corresponding assembly factors of some species are already known, it remains unclear how CENH3 location is determined. Collins et al. (2004) proposed that ubiquitin-mediated proteolysis contributes to restrict CENH3 localization to centromeres in yeast. This is consistent with observations in Drosophila, in which failure of proteolysis resulted in ectopic incorporation of CENH3 (Moreno-Moreno et al. 2006). CENP-A is recruited to DNA double-strand breaks in human and mouse cells lines (Zeitlin et al. 2009). In cell-free egg extracts of *Xenopus laevis*, induced DNA damage increased the number of CENH3 foci to approximately two-fold (Zeitlin et al. 2005). It still remains to be elucidated whether CENH3 has a role in DNA repair.

1.3.3 The role of CENH3 in meiosis

Meiosis is an essential process in the course of germ cell formation in all eukaryotes which reproduce sexually, resulting in gametes or spores. As during mitosis, the kinetochore interacts with the microtubules guiding the meiotic division. However, the function of the kinetochore proteins during the first meiotic division is not completely understood.

Monen et al. (2005) studied the localization of kinetochore proteins during female meiosis in *C. elegans,* which has two highly similar genes encoding CENH3: HCP-3 (Buchwitz et al. 1999) and CPAR-1 (Monen et al. 2005). CPAR-1 is expressed at a very low level, <5% of the level of HCP-3. CeCENP-A, a generic name to refer to both protein isoforms, is absent in early meiotic prophase and was first detected in the late pachytene/diplotene. CeCENP-A signals detected by immunostaining and by imaging of living meiotic embryos decreased during meiosis II in comparison to meiosis I, indicating a removal of CENH3 from chromatin between the two divisions. The selective removal of CENH3 from chromatin during early meiotic prophase is thought to facilitate recombination.

Talbert et al. (2002) had shown CENH3 localization during meiotic division in pollen mother cells (PMC) of *A. thaliana*. During metaphase I up to 10 signals were detected, and arranged as 5 pairs after alignment on the plate.

Because is also essential for meiotic division, CENH3 reduction strongly affects reproduction. Transgenic *A. thaliana* lines expressing RNAi CENH3, revealed formation of micronuclei in pollen tetrads. Besides dwarf phenotype, the plants are partially to almost completely sterile. Transgenic plants EYFP-CENH3(C) carrying only the C-terminal part of CENH3 are also semi to complete sterile, and micronuclei were commonly found in tetrads (Lermontova et al. submitted).

2. Aims of this work

The present work aimed to:

i) Elucidate how much conservation to AtCENH3 is sufficient for targeting of alien CENH3 to *A. thaliana* centromeres;

Previously it was shown that the C-terminal part of AtCENH3, including the loop1 region, is enough to direct the protein to the centromere. Complete AtCENH3 can recognize the three different centromeric repeats of A. lyrata, but can not enter the nuclei of *V. faba*. To investigate how much similarity to the endogenous protein is enough to target alien CENH3 to *A. thaliana* centromeres, *A. thaliana* plants were stably transformed with CENH3 genes cloned from *A. arenosa, A. lyrata, Capsella bursa-pastoris, Zea mays and Luzula nivea*. Plants expressing the alien CENH3 were analyzed for protein localization by immunostaining and life cell imaging.

ii) Find out wether the three most conserved aminoacids at the loop1 region are crucial for CENH3 targeting;

The loop1 region is rather conserved among *Arabidopsis* species, but less so between distantly related plant species. The studied species share only three aminoacids at the loop1 region: threonine2, arginine12 and alanine15. To test their importance for CENH3 targeting, the nucleotides codifying these three aminoacids were changed by mutagenesis PCR by asparagine, proline and valine. Nuclei of plants transformed with the mutated sequence were analyzed by immunostaining for localization of the recombinant AtCENH3.

3. Material and Methods

3.1 Material

3.1.1 Chemicals, enzymes and kits

The chemicals used were obtained from the following companies: Merck (Darmstadt) and Sigma-Aldrich (Steinheim). PCR oligonucleotide primers were synthesized at Metabion (München). Sequencing was performed by Agowa (Berlin) and the IPK sequencing facility. The DNA and protein size markers, as well all the enzymes used, were obtained from Fermentas (St. Leon-Rot). The PCR purification, DNA and RNA extraction and Plasmid isolation kits were obtained from Qiagen (Hilden) and Invitek. Kits for *in situ* mutagenesis PCR were obtained from Finnzymes (Espoo, Finland).

3.1.2 Organisms

Bacteria

For cloning purposes, the following bacteria strains were used:

- *Escherichia coli* Top10 (Invitrogen, Karlsruhe)
- *Agrobacterium tumefaciens* GV3101 (Koncz and Schell 1986)

Plasmids

Plasmid	Selective marker	Reference
p35S-BM	Ampicilin (100µg/µl)	DNA cloning service, Hamburg (Hausmann and Toepfer, 1999)
pLH7000	Spectinomycin (100 µg/µl)	DNA cloning service, Hamburg (Hausmann and Toepfer, 1999)

Plants

The following plant species were used:

- *Arabidopsis thaliana* ecotype columbia (IPK Gatersleben), for plant stable transformation, sorting nuclei, cytological preparations, and RNA and DNA extraction.
- *Luzula nivea* (Benary Samen), for cytological preparations and RNA and DNA extraction.
- *Hordeum vulgare* golden promise variety, for transient expression of constructs.

3.1.3 Media

Plant media

Plant Growth Medium (GM) was used for seed germination:

50 ml macroelements, 500 µl microelements (Ducheva, Haarlem), 10 ml B5 vitamin, 5 ml Fe-EDTA, 500 mg MES, 8 g Agar; Ph 5,7

Bacterial media

For bacteria growing, three different media were used:

- LB, for *E. coli*: 10 g Bacto Tryptone, 5 g Yeast extract, 5 g NaCl, pH 7,0; for solid medium, add 15 g Agar; for liquid medium, add 7 g Agar (Bertani 1951 and 2004)

- SOC, for bacteria incubation after transformation: 20 g Tryptone, 5 g Yeast extract, 0,5 g NaCl, 2,5 mM KCl, 10 mM MgCl2, 20 mM Glucose; pH 7,5 (Hanahan 1983)

- YEB, for *A. tumefaciens*: 5 g Beef extract, 2 g Yeast extract, 5 g Peptone, 5 g Saccharose, MgSO4, 15 g Agar (Miller 1972)

3.1.4 Software

- Oligo Analyzer 3.1

  http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx

  This software was used to analyse the crucial aspects for oligonucleotide designing such as optimal melting temperature and probability of primer dimmer formation.

- NEBCutter 2.0: http://tools.neb.com/NEBcutter2/

  This tool allows to select enzymes as specific cutters for a sequence.

- MultiAlign: http://bioinfo.genotoul.fr/multalin/multalin.html

  Used for alignment of two or more different sequences.

- Adobe photoshop 6.0

  Photoshop was used to adjust colour of images and for merging images of DAPI and immunostaining signals

3.2 Methods

3.2.1 Genomic DNA extraction

Plant genomic DNA was extracted from 100 mg of leaves following Edwards et al. (1991), by grinding the leaves and re-suspending in Extraction Buffer (0.2 M Tris-HCL, pH 9.0; 0.4 LiCl; 25 mM Na2-EDTA and 1% SDS) or using the DNA extraction mini and maxi Kit (Qiagen).

3.2.2 RNA isolation and cDNA synthesis

Total RNA was extracted from 100 mg of grinded flower buds or seedlings using the RNA extraction kit (Invitek and Qiagen) following manufacturer's instructions. RNA concentration was measured on the full-spectrum (220-750 nm) spectrophotometer Nanodrop ND-1000 (Peqlab), and 1µg was used to synthesize the cDNA using the cDNA synthesis kit (invitrogen), according to manufacturer's instructions.

3.2.3 Cloning methods

3.2.3.1 Generation of Constructs

The EYFP DNA sequence was amplified with the primer pair 5'-ACCACTAGTATGGTGAGCAAGGGCGAGGAG-3' and 5'-ACTGGATCCCTTGTACAGCTCGTCCATGCC-3' from the pWEN18 vector, generating an *Spe*I linker sequence at the 5' end and a *Bam*HI linker sequence at the 3' end. The stop codon of EYFP was not included. The amplified fragment was inserted downstream the 35S promoter in the unique *Spe*I and *Bam*HI sites of the p35S-BAM vector.

To generate the p35S:EYFP-CENH3 fusion construct, the CENH3 sequences of *Arabidopsis lyrata*, *Arabidopsis arenosa, Capsella bursa-pastoris, Zea mays* and *Luzula nivea* were amplified from cDNA (Fig. 5 A) with the CENH3-specific primer pairs containing *Xma*I restriction site at the 5' end and a *Sal*I site at the 3' end for cloning into the p35S-BAM-EYFP vector (www.dna-cloning-service.de) vector (table 2). Amplified products and p35S-BAM-EYFP vector sequences were digested with corresponding restriction enzymes and ligated together. Positive colonies were selected on ampicilin containing LB medium. The presence of CENH3 inserts in the 35S-BAM-EYFP vector was confirmed with appropriate restriction enzymes (Fig. 5 B). Selected positive clones were sequenced to check the correctness of CENH3 sequence after PCR amplification and to confirm that CENH3 is inserted in frame with EYFP. The expression cassettes containing 35S CaMV promoter, EYFP-CENH3 fusion variants and Nos terminator sequences were excised, cloned into the binary vector pLH7000 (www.dna-cloning-service.de, Fig. 5 C) containing the phosphinotricine resistance marker via the *Sfi*I restriction site and used to transform *E. coli* by electroporation. Colonies were picked from the selective medium containing spectinomycin and digested to screen for positive clones (Fig. 5 D). Positive clones were used to transform *A. tumefaciens* GV3101. Standard PCR was done to confirm the presence of the construct (Fig. 5 E) in *Agrobacterium tumefaciens* and to screen transformed *A. thaliana* lines. Forward primers for EYFP were designed and used in combination with the CENH3 reverse primer of each species. The PCR mix contained 1 µM of

primers, 2 mM dNTP, 1x reaction buffer, 1U of High Fidelity Taq Polymerase in a final volume of 50 µl. The PCR conditions were: 30 cycles; initial denaturation, 5 min 95ºC; denaturation, 15 sec 95ºC; annealing, 30 sec 50-70ºC; elongation, 45 sec 72ºC; final elongation, 7 min 72ºC.

Table 2 Primers used to amplify CENH3 gene, constructs and for sequencing

Primer	Sequence 5'-3'
35S_f	TGA GAC TTT TCA ACA AAG
35S_r	TCT CCA AAT GAA ATG AAC
Asn_r	Pho-AGG GGC CAA CAT ATG GAT TAT ACT TCT CAC TTC
Pro_Val_f	Pho-CCC CAA ATC AAT CCT TGG ACA GTT GAA GCT CTT GTT
LnCenH3_B190_f	ATA CCC GGG TGG CTC CGC CAC AAA CTC CTA
LnCenH3_A21_f	ATA CCC GGG CCC CCA ATG TTC GCG GCA
LnCenH3A_f	ATGGCTCGGACGAAACACTTCCCCCAATGTTCG
LnCenH3B_f	ATGGCTCGGACGAAACACTTCTCCAATAAAAAG
LnCenH3_f	ATA CCC GGG ATG GCT CGG ACG AAA CAC TTC
LnCenH3_r	TTA TGC ACC TAT TCG CCT AGC
LnCenH3_xmaI_r	CCC GTC GAC TTA TGC ACC TAT TCG CCT AGC
EYFP_speI_f	ACC ACT AGT ATG GTG AGC AAG GGC GAG GAG
Bar_f	AGT CGA CCG TGT ACG TCT CC
Bar_r	TCT GGA TTT TAG TAC TGG ATT TTG G
HvCenH3_f	ATA CCC GGG ATG GCC CGC ACC AAG CAC CCC
HvCenH3_r	CCG GTC GAC TCC TTT TGC ATG ACG GTA ACA
Aly_Aar_CenH3_f	ATA CCC GGG ATG GCG AGA ACC AAG CAT TTC
Cap_CenH3_f	ATA CCC GGG ATG GCG AGA ACA AAG CAT TTC
ZmCenH3_f	ATA CCC GGG ATG GCT CGA ACC AAG CAC CAG
ZmCenH3_r	CCG GTC GAC TCA TGC CCA ACG CCT TCC TCC
AttRI	GGG GAC AAG TTT GTA CAA AAA AGC AGG CT
AttRII	GGG GAC CAC TTT GTA CAA GAA AGC TGG GT
35S_left	GGA GCA CGA CAC TCT GGT CT
35S_right	CGT GTC CTC TCC AAA TGA AA

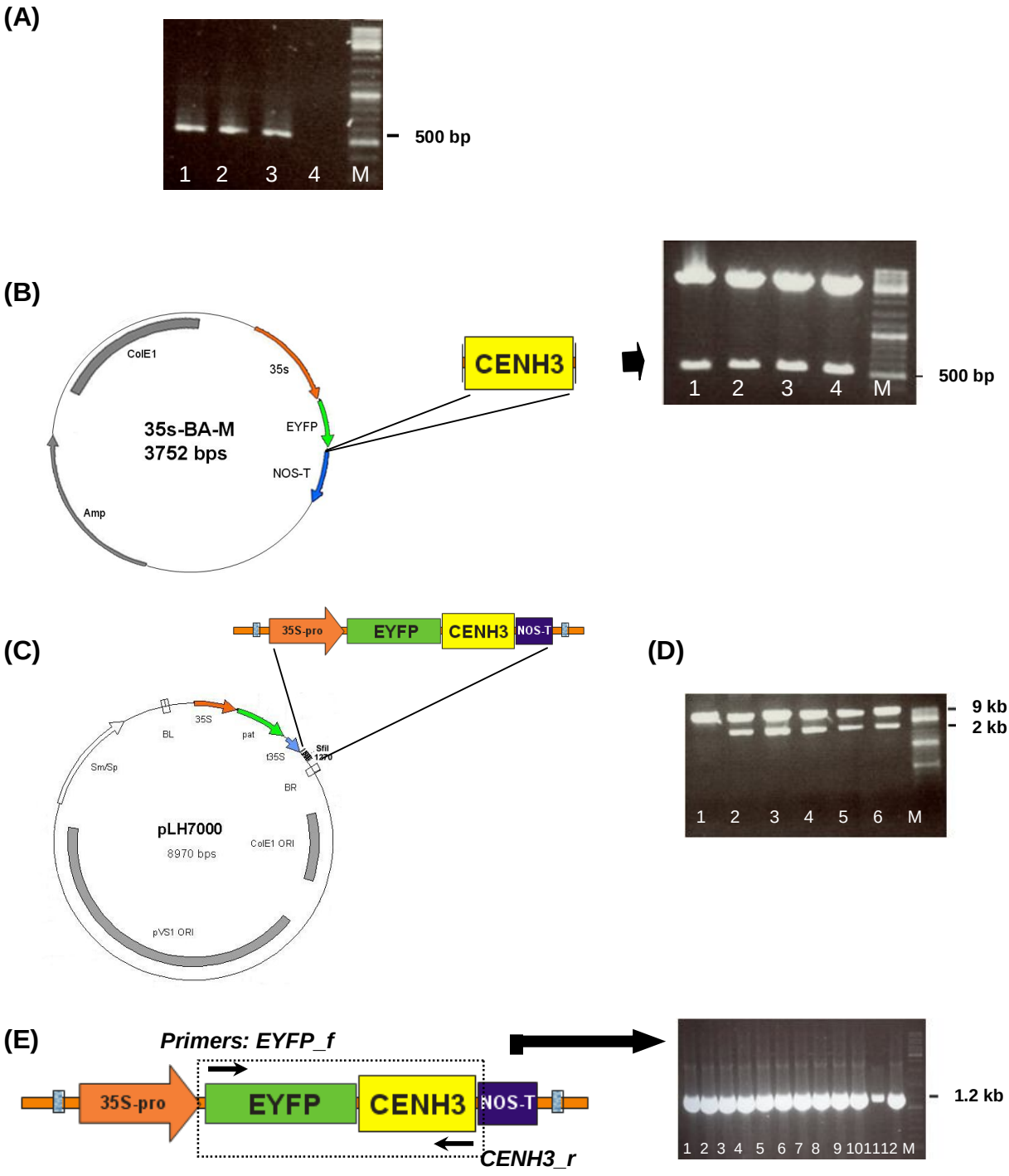

FIGURE 5 Cloning of CENH3 genes **(A)** PCR product of AlyCENH3 amplified *via* PCR, ~ 500bp; **(B)** scheme of the intermediate vector p35-BAM-EYFP, and screening for clones on the gel: positive plasmids are digested into a larger band and one of 500 bp corresponding here to AarCENH3 clones 1-4; **(C)** scheme of the binary vector pLH7000 of 8.97 kb; **(D)** screening for positive clones: gel showing the digested vector with the expression cassette (2 kb) of clones 2-6; **(E)** PCR amplification of EYFP-CENH3 to screen *A. tumefaciens* for plant transformation resulted in a product with the correct size of ~1.2kb.

3.2.3.2 Mutagenesis PCR

To obtain *A. thaliana* plants expressing CENH3 carrying substitutions at the three amino acid positions, mutagenesis PCR with the primer pair Pro_Val_f: 5'-Pho-CCCCAAATCAATCCTTGGACAGTTGAAGCTCTTGTT-3' and Asn_r: 5'-Pho-AGGGGCCAACATATGGATTATACTTCTCACTTC-3'phosphorylated at the 5' was performed using as template the fusion construct p35S::EYFP-AtCENH3. The amplified products and the p35S-BAM-EYFP vector were digested with the enzyme DpnI, which digests methylated sites, eliminating parental plasmids, but preserving the mutated sequence. Selected positive clones were sequenced to confirm the substitutions within the AtCENH3(M) sequence after PCR amplification (Fig. 7). The expression cassettes containing the 35S CaMV promoter, EYFP-CENH3 fusion variants and the Nos terminator were excised, cloned to the binary vector pLH7000 and used to transform *E. coli* by electroporation (Fig. 6 A). Five colonies were picked from the selective medium containing spectinomycin and digested to screen for positive clones. Positive clones were used to transform *A. tumefaciens* GV3101.

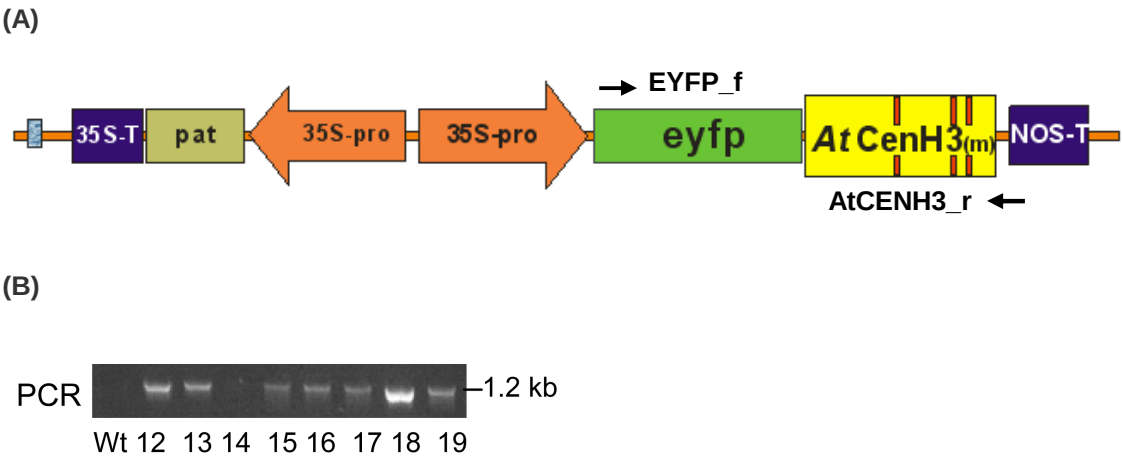

FIGURE 6 Recombinant AtCENH3(M) **(A)** Scheme of the expression cassette containing the mutated CENH3 gene sequence; **(B)** PCR with EYFP_f and AtCENH3_r primers and genomic DNA of transformed lines 12-19 yielded a product of the expected size of 1.2 kb

The colonies of *A. tumefaciens* were screened by PCR amplification with EYFP_f and AtCENH3_r to confirm the presence of the transgene prior to *A. thaliana* transformation. The transformed plants were first screened by selection with the herbicide BASTA, resulting in 28 lines. These lines were selected by PCR amplification with EYFP_f and AtCENH3_r primers (Fig. 6 B) and by immunostaining with anti-GFP on squashed root tips for the expression of recombinant protein. Lines with strong EYFP expression were selected for further analysis.

```
                321
  EYFP-AtCENH3  CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG
EYFP-AtCENH3(M) CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG

                401                                                                             480
  EYFP-AtCENH3  ACGGCAACAT CCTGGGGCAC AAGCTGGAGT ACAACTACAA CAGCCACAAC GTCTATATCA TGGCCGACAA GCAGAAGAAC
EYFP-AtCENH3(M) ACGGCAACAT CCTGGGGCAC AAGCTGGAGT ACAACTACAA CAGCCACAAC GTCTATATCA TGGCCGACAA GCAGAAGAAC

                481                                                                             560
  EYFP-AtCENH3  GGCATCAAGG TGAACTTCAA GATCCGCCAC AACATCGAGG ACGGCAGCGT GCAGCTCGCC GACCACTACC AGCAGAACAC
EYFP-AtCENH3(M) GGCATCAAGG TGAACTTCAA GATCCGCCAC AACATCGAGG ACGGCAGCGT GCAGCTCGCC GACCACTACC AGCAGAACAC

                561                                                                             640
  EYFP-AtCENH3  CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC TACCTGAGCT ACCAGTCCGC CCTGAGCAAA GACCCCAACG
EYFP-AtCENH3(M) CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC TACCTGAGCT ACCAGTCCGC CCTGAGCAAA GACCCCAACG

                641                                                                             720
  EYFP-AtCENH3  AGAAGCGCGA TCACATGGTC CTGCTGGAGT TCGTGACCGC CGCCGGGATC ACTCTCGGCA TGGACGAGCT GTACAAGTAA
EYFP-AtCENH3(M) AGAAGCGCGA TCACATGGTC CTGCTGGAGT TCGTGACCGC CGCCGGGATC ACTCTCGGCA TGGACGAGCT GTACAAGGGA

                721                                                                             800
  EYFP-AtCENH3  ---------A TGGCGAGAAC CAAGCATCGC GTTACCAGGT CACAACCTCG GAATCAAACT GATGCCGCCG GTGCTTCATC
EYFP-AtCENH3(M) TCCCCCGGGA TGGCGAGAAC CAAGCATCGC GTTACCAGGT CACAACCTCG GAATCAAACT GATGCCGCCG GTGCTTCATC

                801                                                                             880
  EYFP-AtCENH3  TTCTCAGGCG GCAGGTCCAA CTACGACCCC GACAAGGAGA GGCGGTGAAG GTGGAGATAA TACTCAACAA ACAAATCCTA
EYFP-AtCENH3(M) TTCTCAGGCG GCAGGTCCAA CTACGACCCC GACAAGGAGA GGCGGTGAAG GTGGAGATAA TACTCAACAA ACAAATCCTA

                881                                                                             960
  EYFP-AtCENH3  CAACTTCACC AGCTACTGGT ACAAGGAGAG GGGCTAAGAG ATCCAGACAG GCTATGCCAC GAGGCTCACA GAAGAAGTCT
EYFP-AtCENH3(M) CAACTTCACC AGCTACTGGT ACAAGGAGAG GGGCTAAGAG ATCCAGACAG GCTATGCCAC GAGGCTCACA GAAGAAGTCT

                961                                                                            1040
  EYFP-AtCENH3  TATCGATACA GGCCAGGAAC CGTTGCTCTA AAAGAGATTC GCCATTTCCA GAAGCAGACA AACCTTCTTA TTCCGGCTGC
EYFP-AtCENH3(M) TATCGATACA GGCCAGGAAC CGTTGCTCTA AAAGAGATTC GCCATTTCCA GAAGCAGACA AACCTTCTTA TTCCGGCTGC

               1041                                                                            1120
  EYFP-AtCENH3  CAGTTTCATA AGAGAAGTGA GAAGTATAAC CCATATGTTG GCCCCTCCCC AAATCAATCG TTGGACAGCT GAAGCTCTTG
EYFP-AtCENH3(M) CAGTTTCATA AGAGAAGTGA GAAGTATAAT CCATATGTTG GCCCCTCCCC AAATCAATCC TTGGACAGTT GAAGCTCTTG

               1121                                                                            1200
  EYFP-AtCENH3  TTGCTCTTCA AGAGGCGGCA GAAGATTACT TGGTTGGTTT GTTCTCAGAT TCAATGCTCT GTGCTATCCA TGCAAGACGT
EYFP-AtCENH3(M) TTGCTCTTCA AGAGGCGGCA GAAGATTACT TGGTTGGTTT GTTCTCAGAT TCAATGCTCT GTGCTATCCA TGCAAGACGT

               1201                                                            1266
  EYFP-AtCENH3  GTTACTCTAA TGAGAAAAGA CTTTGAACTT GCACGCCGGC TTGGAGGAAA AGGCAGACCA TGGTGA
EYFP-AtCENH3(M) GTTACTCTAA TGAGAAAAGA CTT-GAAC
```

FIGURE 7 Sequence alignment between one of the selected clones carrying the original (EYFP-AtCENH3) and mutated (EYFP-AtCENH3[M]) nucleotides (box).

3.2.3.3 Semi-quantitative Reverse Transcriptase PCR (sqRT-PCR)

Semi-quantitative RT-PCR was perfomed using the primers described on table 1, with the following temperatures: 95ºC for 10 sec, 58ºC for 45 sec and 72ºC for 90 sec. The samples (5 µl) were collected and loaded on a 1% agarose gel after 17, 22, 27, 32 and 37 cycles, respectively.

3.2.3.4 Plasmid preparation

All plasmid DNAs were prepared using the QIAprep kit (Qiagen, Hilden), according to the supplier´s instructions.

3.2.3.5 Electroporation

Competent cells were mixed with 4 µl of the ligation mixture and left for 1 h at room temperature, then filtered through a nitrocellulose membrane for 30 min. The transformation mix was transferred into a cuvette and submitted to an electric pulse using the Gene Pulser II Electroporation system from BioRad (BioRad, USA) at the following settings: voltage of 2.5 kV; capacitance of 25 µF; impedance at low Ohm; Voltage Booster resistance at 200 Ohm; impulse for ca. 5 msec.

After electroporation 450 µl of SOC medium was added to the cuvette, mixed and transferred to a 2 ml tube. The bacterial suspension was incubated under shaking at 200 rpm for 1 h at 37°C in the case of *E. coli* or 28ºC in the case of *A. tumefaciens*. About 50-100 µl of the bacterial suspension were plated on LB or YEB plates containing ampicillin (200 µg/ml), spectinomycin (60 µg/ml) and rifampicin (60 µg/ml).

3.2.3.6 Stable and transient plant transformation

The 35S:EYFP-CENH3 plasmid was tested for expression by gold particle bombardment in *Hordeum vulgare* leaves (Schenk et al. 1998). Transient transformation was performed with 5 mg of plasmid DNA on 25-mg gold particles using a Biolistic PDS-1000/HE system according to manufacturer's instructions (Bio-Rad). After bombardment, the leaves were incubated for 16 h under light/dark conditions prior to EYFP localization *via* confocal microscopy.

Plants of *A. thaliana* accession Columbia were transformed according to the flower-dip method (Bechtold et al. 1993). Transgenic EYFP-CENH3–containing progenies were selected after surface sterilization of seeds on Murashige and Skoog medium (Murashige and Skoog 1962) containing 8 mg/L of phosphinotricine. Growth conditions in a cultivation room were 20°C 16 h light/18°C 8 h dark.

3.2.4 Protein studies

3.2.4.1 Protein extraction and protein expression

Plant material (100 mg) was frozen, grinded and suspended in 500 µL of solubilization buffer (56 mM Na_2CO_3, 56 mM dithiothreitol, 2% SDS, 12% sucrose, and 2 mM EDTA). After 15 min of incubation at 70°C, the remaining tissue debris were removed by centrifugation. Protein concentration was determined according to Bradford (1976).

3.2.4.2 Western Blotting

Protein samples were separated by SDS-PAGE in 12.5% polyacrylamide gels according to Laemmli (1970). After blotting, membranes were incubated for 12 h at 10°C in TBST and 4% low-fat milk containing anti-CENH3, anti-GFP, or the secondary antibodies. Secondary anti-rabbit (Bio-Rad) antibodies conjugated to horseradish peroxidase were used to visualize immunocomplexes by an enhanced chemiluminescence detection kit (Bio-Rad) according to manufacturer's instructions.

3.2.5 Cytological studies

3.2.5.1 Isolation and Flow Sorting of Nuclei

For the isolation of leaf nuclei, plants were grown in soil in a cultivation room. Leaves were fixed in 4% formaldehyde in Tris buffer (10 mM Tris, 10 mM Na_2EDTA, and 100 mM Triton X-100, pH 7.5) for 20 min. Nuclei were isolated, stained with DAPI (1 µg/mL), and processed for flow sorting according to their fluorescence intensity reflecting the DNA content as described (Jasencakova et

al. 2000). Approximately 1000 nuclei of each fraction were sorted onto microscopy slides into a drop containing 100 mM Tris, 50 mM KCl, 2 mM $MgCl_2$, 0.05% Tween 20, and 5% sucrose, air-dried, and used for immunolabeling or stored at 4°C.

3.2.5.2 Chromosome Preparation

Seeds of the WT and of transformed *A. thaliana* were germinated in Petri dishes on wet filter paper for 3 days at room temperature. Seedlings were fixed for 20 min with ice-cold 4% (w/v) paraformaldehyde in MTSB buffer (50 mM PIPES, 5 mM $MgSO_4$, and 5 mM EGTA, pH 6.9). After washing 3 x 5 min in MTSB, the seedlings were digested at 37°C for 10 min with a PCP enzyme mixture (2.5% pectinase, 2.5% cellulase Onozuka R-10, and 2.5% Pectolyase Y-23 [w/v] dissolved in MTSB). After washing 2 x 5 min in MTSB, root tips were squashed in a drop of MTSB buffer. After freezing in liquid nitrogen, the cover slips were removed and the slides were immediately transferred into MTSB.

3.2.5.3 Immunostaining

Immunostaining of nuclei/chromosomes was performed as described (Jasencakova et al. 2000). EYFP-CENH3 was detected with rabbit polyclonal antisera against GFP (1:500; BD Biosciences) and goat anti-rabbit Alexa 488 (1:200; Sigma-Aldrich). Endogenous CENH3 was detected using antibodies against *A. thaliana* CENH3 (1:500) and goat anti-rabbit rhodamine (1:200; Jackson Immuno Research Laboratories).

LnCENH3-B was detected using guinea-pig anti-LnCENH3 (1:100, Covalab) and anti-guinea pig alexa 488 secondary antibodies.

The guinea-pig antibody against LnCENH3-B, generated by Covalab Ltd, Cambridge, was affinity-purified against peptides representing amino acids 8-20 and 61-72 of LnCENH3-B.

3.2.5.4 Confocal Microscopy Analysis

A. thaliana seeds of lines harbouring EYFP-ZmCENH3, EYFP-CpbCENH3 and EYFP-LnCENH3 were germinated in agar medium in cover slip chambers (Nalge Nunc Int.). Roots growing parallel to the cover slip bottom were analysed

in a LSM 510META confocal microscope (Carl Zeiss, Germany) using a 63x oil immersion objective (n.A. 1.4). EYFP was excited with a 488 nm laser line and fluorescence recorded with a 505-550 bandpass filter. Images were analyzed with the LSM software release 3.2.

3.2.5.5 Alexander Staining of pollen grains

Flowers and flower buds were collected in 10% ethanol and incubated overnight at 10°C. Anthers were isolated and put on slides. Dissected anthers were incubated with Alexander stain (Alexander 1969) under coverslips for 15 min at room temperature and evaluated using a light microscope Axiophot (Carl Zeiss, Germany).

4. Results

4.1. Phylogenetic analysis of CENH3 of various plant species to identify their degree of similarity to AtCENH3

To understand how much similarity to AtCENH3 is needed for *A. thaliana* centromere targeting by alien CENH3, sequences of closely and more distantly related species were selected to transform *A. thaliana* WT and to analyse the sub cellular localization of the alien protein. Except for L. *nivea*, all the selected plant species have monocentric chromosomes.

Although initially the aim was to study T-DNA insertion lines, public databases did not list any CENH3 *A. thaliana* mutant line. Because CENH3 is an essential protein, a complete loss of function mutation might be lethal. Consequently, *A. thaliana* WT were used to generate lines of transformants instead.

The sequence similarity of plant CENH3 protein was determined by alignment according to the Jotun Hein method (Hein 1990). The relatedness of these CENH3 sequences is displayed in a phylogenetic tree (Fig. 8). AtCENH3 shares a high similarity with closely related species, decreasing with the taxonomic distance.

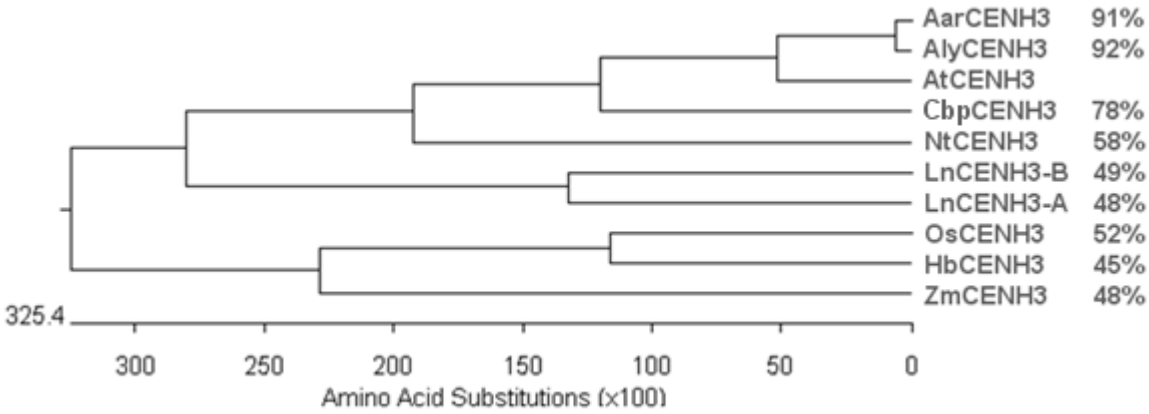

FIGURE 8 Phylogenetic tree of CENH3 protein of plant species aligned according to the Jotun Hein method with the percentage of similarity of the studied plant species to *A. thaliana (*AtCENH3). Aar = *A. arenosa*; Aly = *A. lyrata*; Cbp = *Capsella bursa-pastoris*; Nt = *Nicotiana tabacum*; Ln = *Luzula nivea*; Os = *Oryza sativa*; Hb = *Hordeum bulbosum*; Zm = *Zea mays*.

Sequence analyze of the cloned CENH3 genes confirmed the correct CENH3 sequence for all studied plant species and revealed two distinct sequences of *A. lyrata*. As described by Kawabe et al. (2006), there are two CENH3 genes (called HTR12A and HTR12B) in this species. HTR12B is highly polymorphic between subspecies, while HTR12A shows only few polymorphisms. The many polymorphisms found in the sequence of the cloned AlyCENH3 gene indicate that is the isoform B.

4.1.1 Identification of two active CENH3 genes in L. nivea

In the course of efforts to clone the LnCENH3 gene, the PCR products amplified from cDNA revealed two distinct sequences (A and B): One (A) corresponding to the gene described by Nagaki et al. (2005); the second one (B) shared the first 21 and last 37 nucleotides, but displayed many nucleotide substitutions and a 31 bp insertion in the remaining part (Fig. 9).

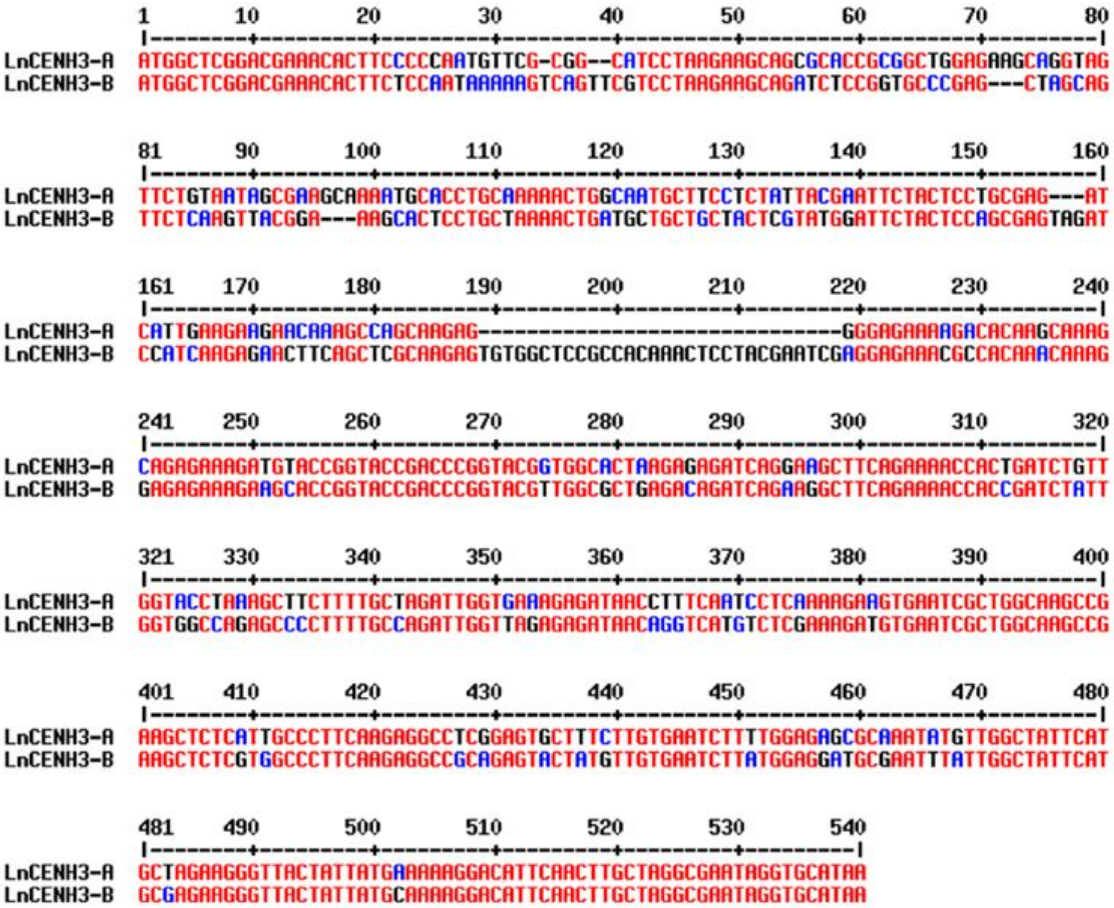

FIGURE 9 Nucleotide sequence alignment between the two LnCENH3 isoforms isolated from cDNA

A BlastN search did not yield a match compatible with the second sequence. The sequence was submitted to GenBank and is currently available under the accession number HM988988.

In order to confirm the presence of two isoforms of *L. nivea*, both LnCENH3-A and -B sequences were submitted to Nebcutter and enzymes were selected to digest A and B at specific sites. Restriction digestion of cloned cDNA confirmed the existence of the two isoforms. The enzyme SacII cuts LnCENH3-A at nucleotide position 56, resulting in bands of 56 and 447 bp. LnCENH3-B is digested by ScaI into bands of 430 and 104 bp (Fig. 10). The additional larger band represents the undigested second isoform of 534 (LnCENH3-B) and 503 bp (LnCENH3-A), respectively.

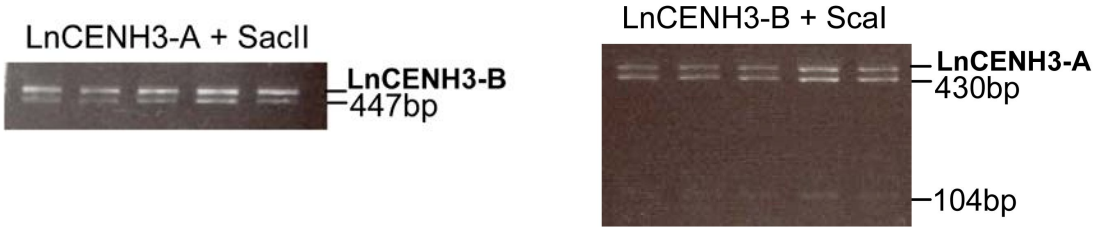

FIGURE 10 Restriction digestion confirmed the existence of two CenH3 genes: SacII cuts the isoform A in two bands of 447bp and 56bp and ScaI cuts isoform B in two bands of 430bp and 104bp respectively. The larger bands on both gels represent the undigested isoform.

LnCENH3-A and -B genes were amplified from cDNA showing that both are transcribed. To test whether LnCENH3-B protein is active and where it is located within the nuclei and on chromosomes, two peptides present on LnCENH3-B (Fig. 11) but absent in LnCENH3-A were used to raise the specific antibody (Covalab) in guinea-pig.

FIGURE 11 Peptides absent on LnCENH3-A were chosen to synthesize a LnCENH3-B specific antibody.

Immunostaining on *L. nivea* root tip nuclei with affinity-purified anti-LnCENH3-B (Fig. 12 A-B) revealed the same distribution pattern as described by Nagaki et al. (2006) for LnCENH3-A. The signals are visible after late prophase, but not during interphase.

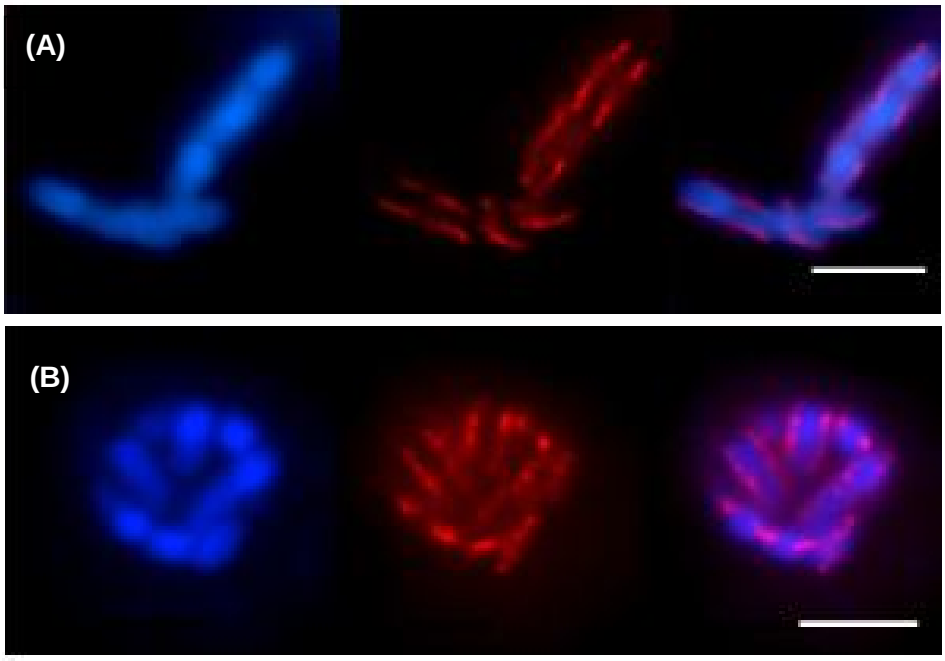

FIGURE 12 Immunostaining experiments confirmed the expression and chromosomal localization of LnCENH3-B. **(A-B)** metaphase chromosomes. Blue = DAPI; red = anti-LnCenH3-B; bar = 5µm

4.1.2 Transient expression of EYFP-alienCENH3 in barley leaves confirmed the expression of the constructs

In order to test the expression of the prepared constructs prior to stable transformation of *A. thaliana* WT plants, selected clones were transiently expressed in *Hordeum vulgare* leaves *via* gold particle bombardment. After bombarded, the plates containing barley leaves were kept under long day conditions for 2 days prior to analysis. All the constructs carrying EYFP-CENH3 displayed fluorescence signals, confirming their translation.

Two additional CENH3 clones were included in this experiment: HvCENH3, as a positive control, and LnCENH3-A. Although this experiment was carried on only to test the constructs for translation, an obvious variation regarding the distribution of signals was observed in the bombarded leaves: i) the fluorescence signals of alien protein localized at the chromocenters, and/or ii) dispersed signals were distributed over the nuclei.

For AarCENH3, CapCENH3 and ZmCENH3 both patterns were observed, while AlyCENH3, AtCENH3, LnCENH3 A and B and even the control HvCENH3 yielded only dispersed signals (Fig. 13, Table 3).

As other plants, barley deposits CENH3 mainly in late G2 (table 1). Although the youngest leaves were selected for bombardment, it is possible that most of the bombarded cells were already differentiated. Therefore, the inability of incorporation of CENH3 into the centromeres results in dispersion of the recombinant protein throughout the nuclei.

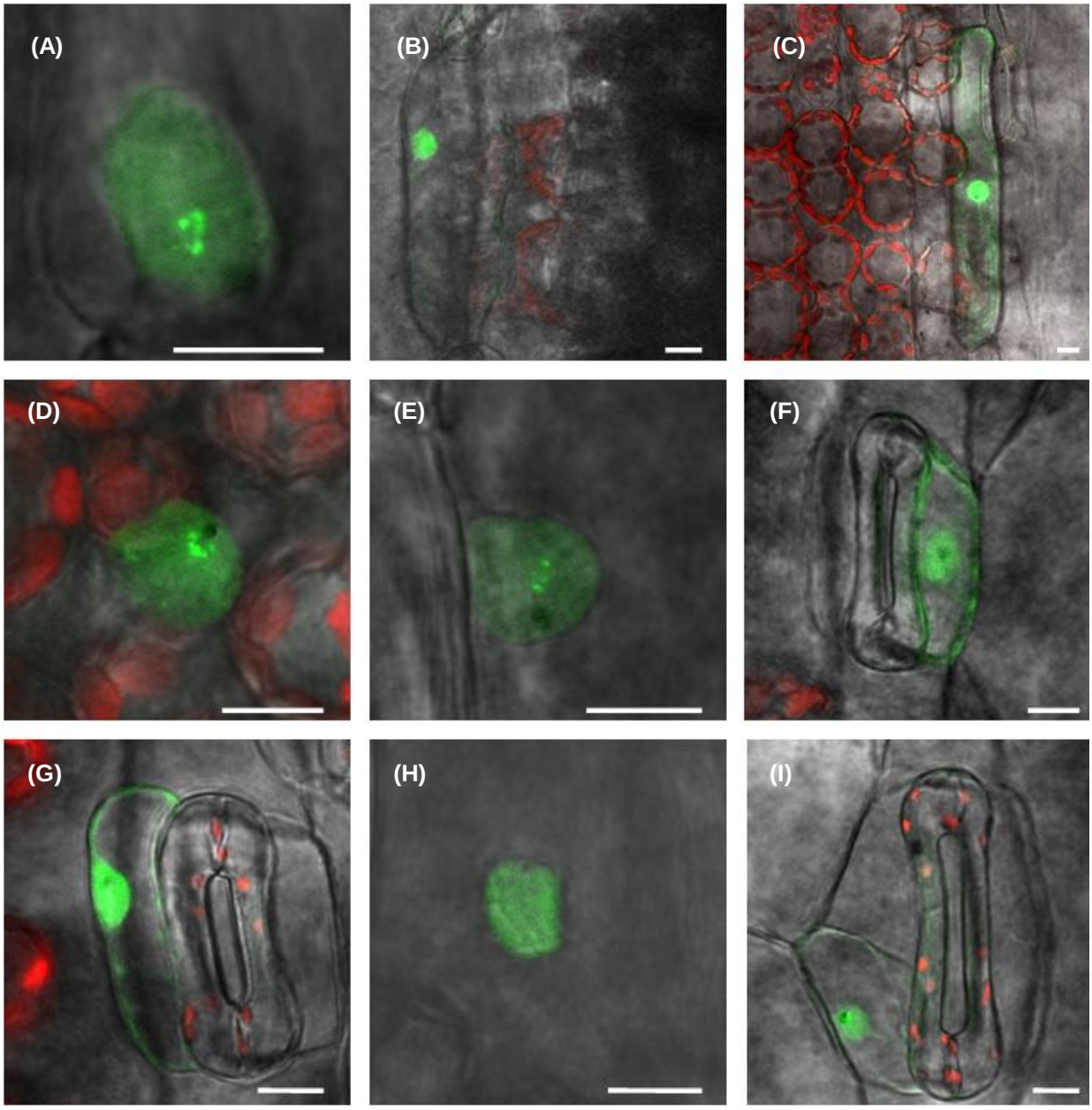

FIGURE 13 EYFP-CENH3 (green) transiently expressed in *Hordeum vulgare* leaves revealed translation of the constructs: AarCENH3 **(A)**, AlyCENH3 **(B)**, AtCENH3**(C)**, CapCENH3 **(D)**, ZmCENH3 **(E)**, HvCENH3**(F, G)**, LnCENH3-A **(H)** and LnCENH3-B **(I)**. AarCENH3, CbpCENH3 and ZmCENH3 display different localization: at the chromocenters (A, B, G, H, I), while the other constructs localized homogeneously on the nucleoplasm. Red = cloroplasts; bar = 10µm

4.1.3 Different alien CENH3 proteins displayed different localization patterns in *A. thaliana* nuclei

After transformation of *A. thaliana* WT plants, the seeds were collected and selected for resistance to the herbicide BASTA. This first screen yielded 9 clones transgenic for AarCENH3, 5 for AlyCENH3, 10 for CbpCENH3, 13 for ZmCENH3 and 9 for LnCENH3-B. These resistant lines were screened for EYFP-CENH3 expression under the fluorescence microscope, but due to fast bleaching, EYFP fluorescence was not reliably detectable. Immunostaining of sorted nuclei or squashed root tips with anti-GFP antibodies turned out to be a suitable alternative to select the lines expressing the recombinant protein.

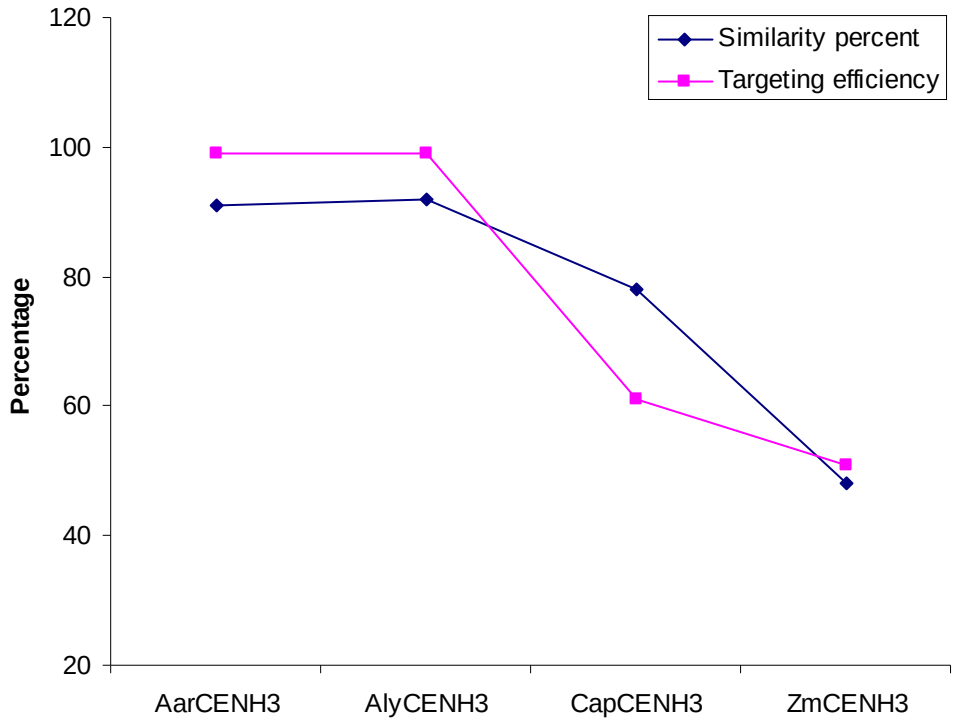

FIGURE 14 Correlation of similarity percent of alien CENH3 to endogenous and their respective targeting efficiency to *A. thaliana* centromeres.

Table 3 Nuclear localization of alien CENH3 in stably transformed lines

Construct	Subcellular localization
AarCENH3	Centromeric
AlyCENH3	Centromeric
AtCENH3	Centromeric
CbpCENH3	Cen / Diffuse / Cen + Diffuse
ZmCENH3	Cen / Diffuse / Cen + Diffuse
LnCENH3-B	No fluorescence signals within the nucleus

For each alien CENH3 construct, all the selected lines were analyzed. The different patterns observed for the nuclear localization of the recombinant proteins is shown in the table 3. There is an obvious correlation of percent of similarity with targeting efficiency (Fig. 14), indicating that the closer related the protein is, the higher is the targeting efficiency. As LnCENH3 was unable to enter the nuclei, it is not possible to estimate targeting efficiency to A. thaliana centromeres, therefore this data was not included in the graphic.

4.1.4 *A. arenosa* and *A. lyrata* CENH3 are targeted to *A. thaliana* centromeres similarly as AtCENH3

Immunostaining with anti-GFP on squashed root tips revealed that AarCENH3 (Fig. 15) and AlyCENH3 (Fig. 16) could target *A. thaliana* centromeres, similarly as does AtCENH3, in all transgenic plants. Most nuclei revealed strong signals at centromeres. Few nuclei displayed diffuse signals.

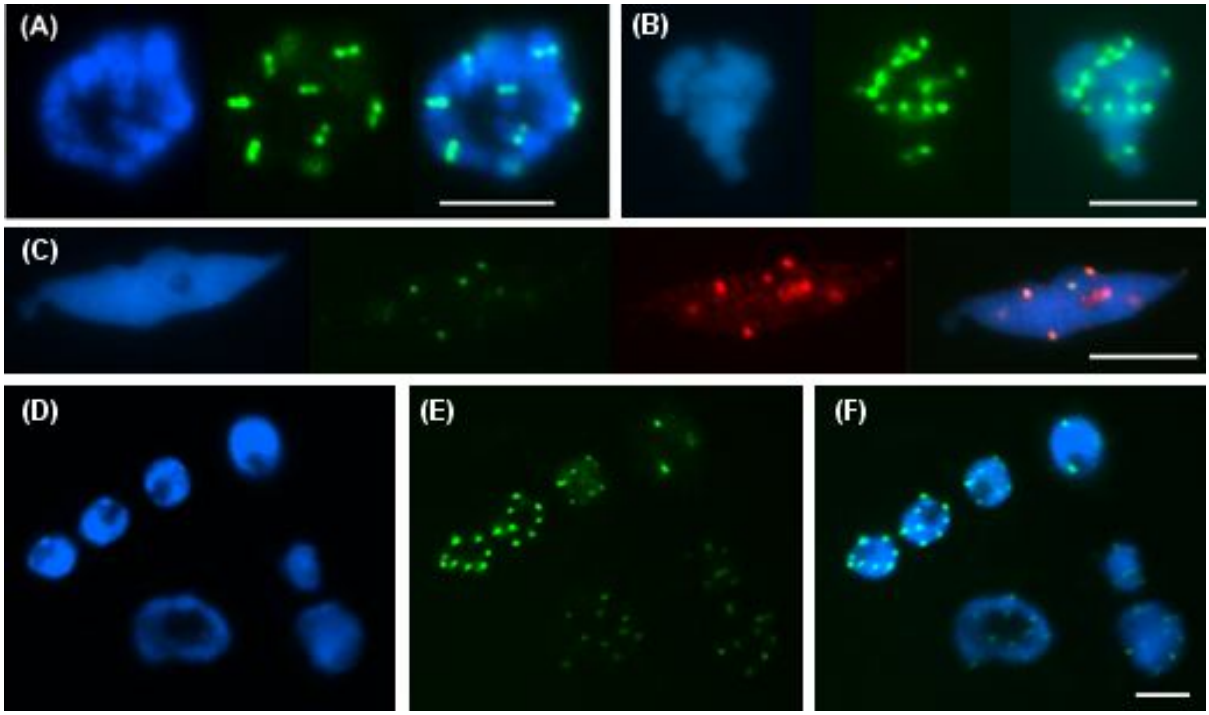

FIGURE 15 Immunostaining of root tips of *A. thaliana* lines expressing AarCENH3: **(A)** G2 nucleus; **(B)** metaphase nucleus; **(C)** double immunostaining with anti-CENH3 and anti-GFP reveals co-localization; **(D), (E)** and **(F)** nuclei showing EYFP-CENH3 signals of different intensity. Blue = DAPI; red = anti-AtCenH3; green = anti-GFP; bar = 5µm

The double signals in G2 and metaphase nuclei indicate that loading of the recombinant protein occurs according to the pattern described for the endogenous AtCENH3 (Lermontova et al. 2006). Double staining with anti-CENH3 and anti-GFP showed co-localization of the recombinant and the endogenous CENH3 (Fig. 15 C).

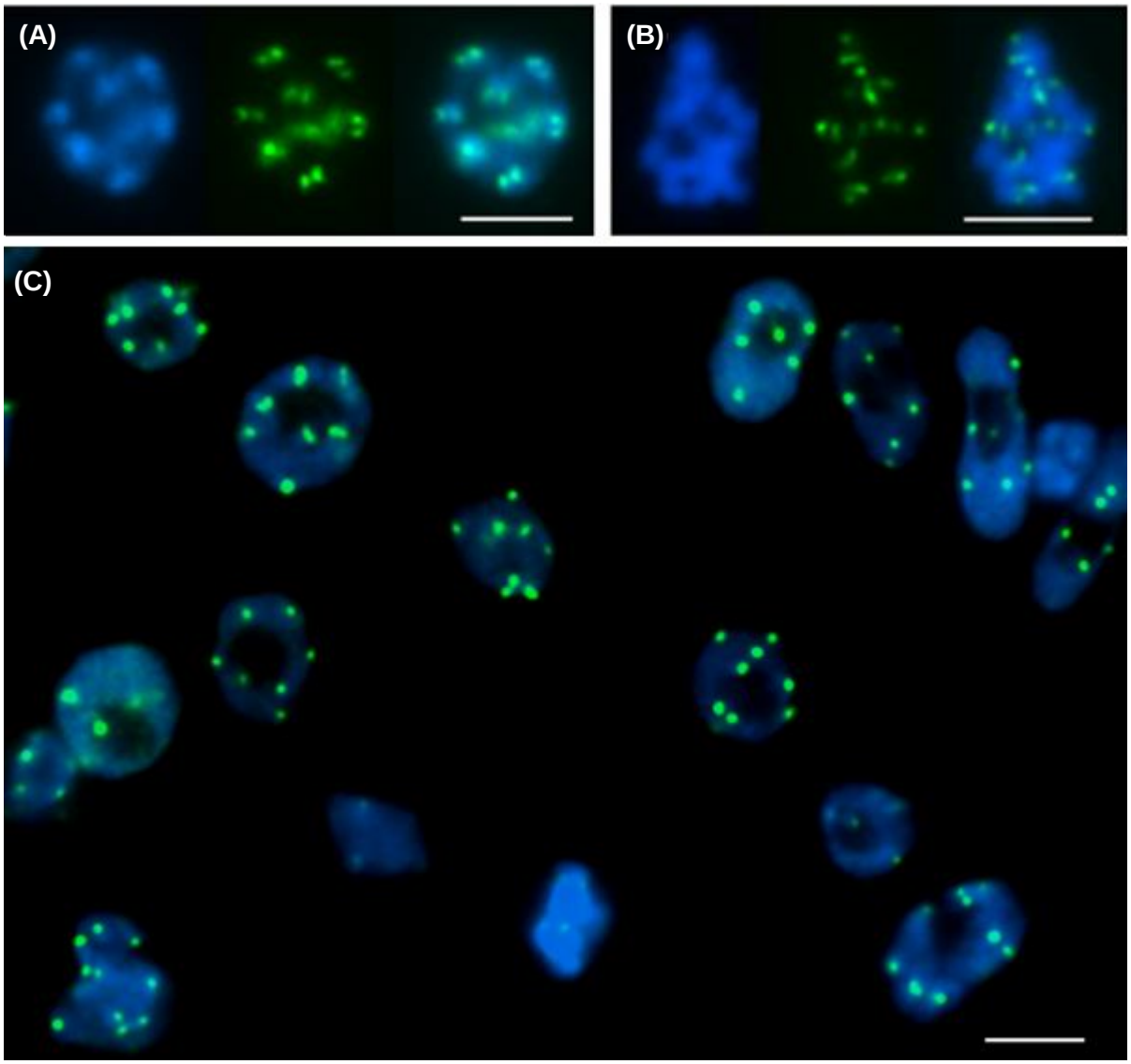

FIGURE 16 Immunostaining of root tips of *A. thaliana* lines expressing AlyCENH3: **(A)** G2 nucleus; **(B)** metaphase nucleus; **(C)** nuclei showing EYFP-CENH3 signals of different intensity. Blue = DAPI; green = anti-GFP; bar = 5µm

4.1.5 C. bursa-pastoris and *Zea mays* CENH3 are less efficient in targeting *A. thaliana* centromeres

CbpCENH3 (Fig. 17) and ZmCENH3 (Fig. 18) were also able to target A. thaliana centromeres. However, lines expressing these alien proteins revealed three different localization patterns, in contrast to AlyCENH3 and AarCENH3, which were almost always able to target *A. thaliana* centromores.

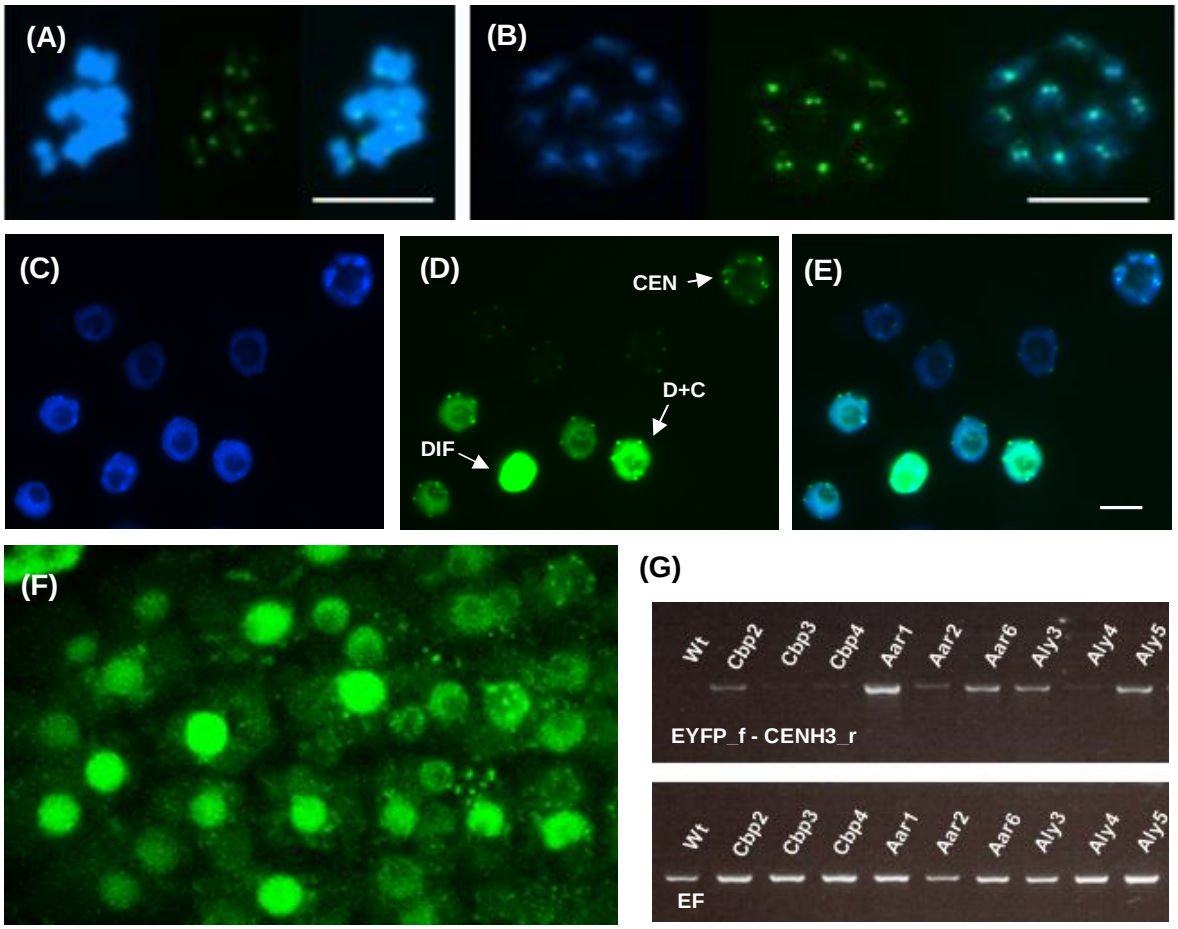

FIGURE 17 Analysis of *A. thaliana* lines expressing CbpCENH3: **(A)** metaphase; **(B)** G2 nucleus; **(C-E)** immunosignals of three different patterns: centromeric localization (CEN), diffuse over the nucleoplasm (DIF) and combination of both (D+C) and **(F)** life cell imaging under confocal microscopy confirms existence of three patterns. Blue = DAPI; green = anti-GFP; bar = 5µm ; **(G)** RT-PCR revealed transcripts after 22 cycles and show the three patterns of localization did not result from over expression; EF = elongation factor gene used as a control.

A considerable percentage of cells displayed diffuse EYFP-CbpCENH3 and EYFP-ZmCENH3 signals over the nuclei and some nuclei showed simultaneously both patterns (table 4). Double immunostaining with anti-GFP and anti-HTR12 antibodies yielded co-localisation of signals for endogenous and recombinant ZmCENH3 within centromeric chromocenters (Fig. 18 A).

The three different localization patterns were common in all the 10 lines expressing CbpCENH3 and the 13 lines transgenic for ZmCENH3 with some

variation of the percentage of each pattern between lines. In total 500 nuclei of 4 lines were counted.

Table 4 Nuclear localization of alien CENH3 in *A. thaliana* nuclei

Localization	CbpCenH3	ZmCenH3
Centromeric	61% ± 6	51% ± 0.5
Centromeric + diffuse	24% ± 5	24% ± 5
Diffuse	15% + 0.5	20% + 1

Lermontova et al. (submitted) observed three different patterns of endogenous CENH3 localization in transgenic *A. thaliana* RNAi lines for CENH3 and in WT young and old leaves. However, these patterns occurred in leaves apparently because CENH3 levels at chromocenters decrease with the age. In root tips, where there is an intense mitotic activity, non-centromeric diffuse signals are very rare.

To exclude that the homogeneous distribution of the fluorescence signals are an artifact of immunostaining, 14-day old seedlings of lines expressing EYFP-CapCENH3 (Fig.17 F) and EYFP-ZmCENH3 (Fig. 18 E) were analyzed under the confocal microscope for EYFP fluorescence. This experiment confirmed the existence of three patterns.

To test whether the diffuse fluorescence signals resulted from over expression, cDNA of some of these lines was used for a RT-PCR reaction. Because the transformed lines expressing AlyCENH3 and AarCENH3 had mostly centromeric localization, some of them were included as a control. Differences at transcription level were observed between lines carrying the same transgene, but no obvious increase on amplified products of CapCENH3 was detected (Fig. 17 G). Therefore, the three different expression patterns do not result from over expression.

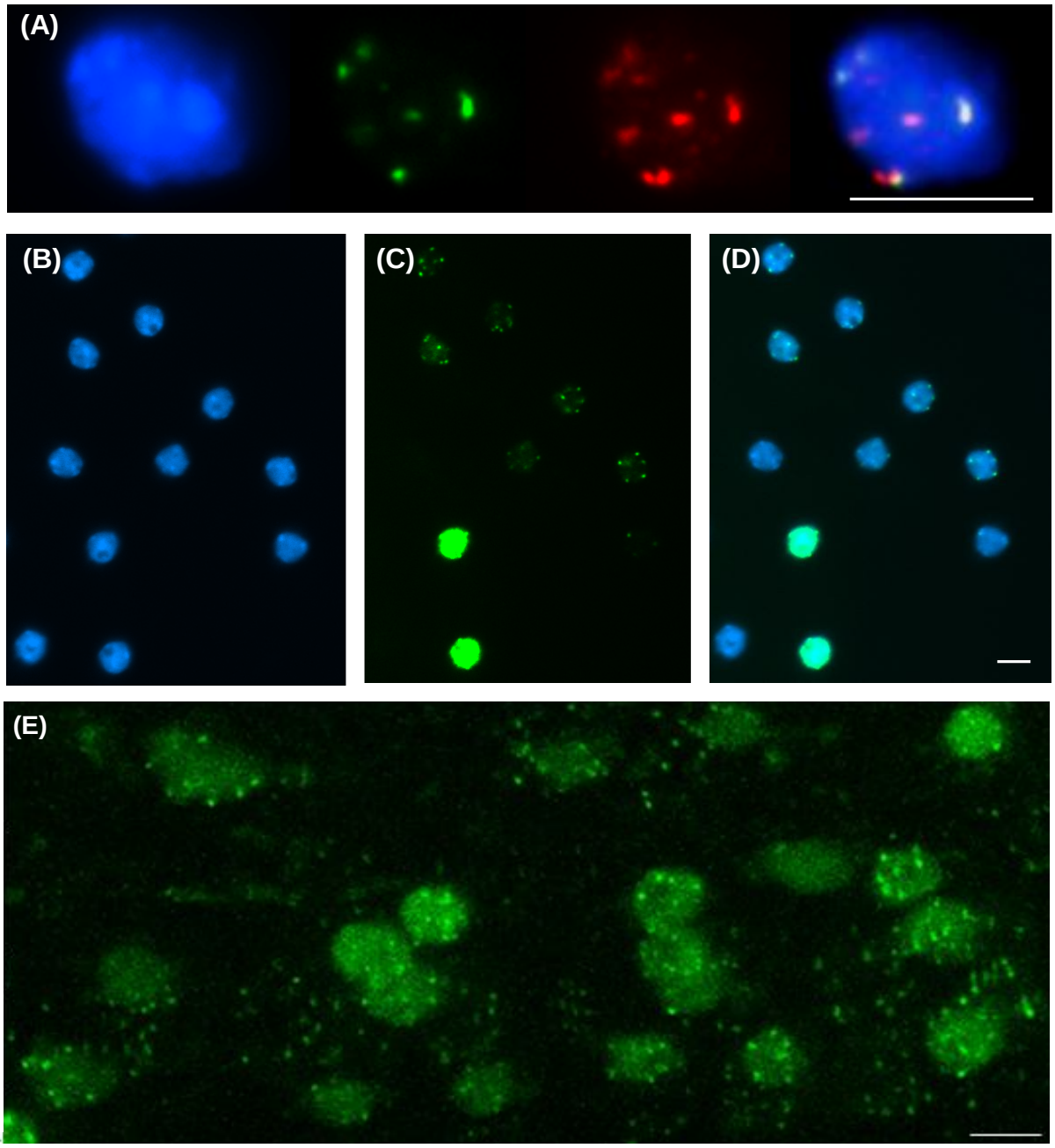

FIGURE 18 Immunostaining and in vivo analysis of root tip nuclei of *A. thaliana* lines expressing ZmCENH3: **(A)** co-localization of endogenous (red) and recombinant (green) protein; **(B-D)** three different patterns of localization were also found on nuclei carrying ZmCENH3. Blue = DAPI; red = anti-AtCenH3; green = anti-GFP; bar = 5 µm **(E)** life cell imaging under confocal microscopy confirms existence of three patterns; bar = 10 µm.

4.1.6 *L. nivea* CENH3 was detectable in cytoplasm but not in nuclei of *A. thaliana*

For the 12 selected lines carrying LnCENH3-B, immunostaining yielded no fluorescence signals (Fig. 19 A-B). However, PCR with EYFP_f and LnCENH3_r primers using *A. thaliana* genomic DNA from EYFP-LnCENH3 transformants confirmed the presence of the construct.

To test whether the transgene is transcribed, RT-PCR was done using cDNA isolated from young leaves, resulting in a transcript of the expected size of 1.2 kb (Fig. 19 E). In order to investigate whether the lack of fluorescence is caused by degradation of the recombinant protein, western blotting with anti-GFP antibodies against protein extracts of EYFP-LnCENH3-B transformants was done. As controls for RT-PCR and for western blotting, lines expressing AlyCENH3, AarCENH3, and CapCENH3, which localized at *A. thaliana* centromeres, were included. Although the expected molecular weight of EYFP-CENH3 is 45 kDa, all recombinant proteins migrated as ~36 kDa (Fig. 19 F). However, sequencing analysis revealed correct sequences of fusion products without stop codons in between, RT-PCR yielded correct size of transcript, and the recombinant proteins (with exception of LnCENH3) had mostly the expected sub cellular localization. Thus, the unexpected molecular weight of EYFP-CENH3 might be due to an anomalous migration, very common for histones.

Since the western blot suggests that LnCENH3 is not degraded, the intact protein is apparently unable to enter the nucleus. To test if EYFP-LnCENH3 localized outside the nucleus, roots of 4-days old seedlings were analyzed on confocal microscope for EYFP fluorescence. The fluorescence signals were distributed over the cytoplasm (Fig. 19 C). Some of the nuclei showed strong green signals possibly because they are embedded in cytoplasm.

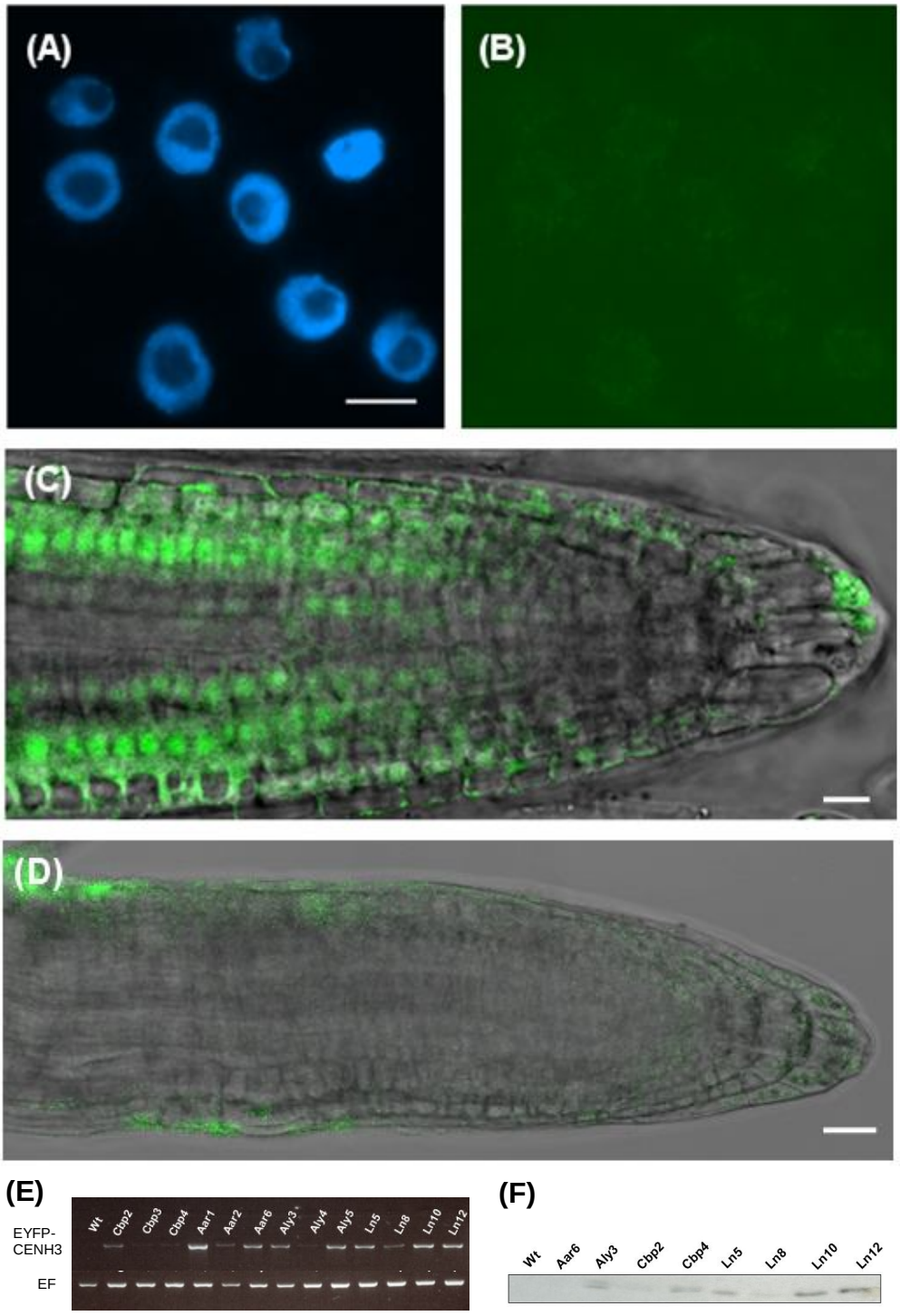

FIGURE 19 LnCENH3 is homogeneously distributed in the cytoplasm but does not enter the nuclei of *A. thaliana:* **(A)** and **(B)** Absence of nuclear localization of LnCenH3-B in *A. thaliana* squashed root tips after immunostaining; bar = 5µm **(C)** root tip revealed EYFP fluorescence over the cytoplasm of LnCENH3 line 5 and **(D)** absence of signals on the negative control, *A. thaliana* WT; bar = 10µm **(E)** RT-PCR confirms transcription of EYFP-LnCENH3 and **(F)** Western blotting with anti-GFP reveals translation of the recombinant protein.

4.2 Three conserved aminoacids (T/2, R/12, A/15) of the loop1 region of CENH3 were substituted (N/2, P/12, V/15) and the mutated construct was transformed into *A. thaliana*

By alignment analysis, three aminoacids at the loop1 region were identified as conserved among the plant species used for this work, as shown below.

```
 AtCenH3 SFIREVRSIT HMLAPPQINR WTAEALVALQ
AarCenH3 SFIRQVRSIT HALAPPQINR WTAEALVALQ
AlyCenH3 SFIRQVRSIT HALAPPQINR WTAEALVALQ
CapCenH3 AFIRQVRSIT NAVAPREVNR WTAEALVALQ
LnCenH3A SFARLVKEIT FQSS-KEVNR WQAEALIALQ
LnCenH3B PFARLVREIT GHVS-KDVNR WQAEALVALQ
 ZmCenH3 PFVRVVRELT NFVTNGKVER YTAEALLALQ
```

The presence of the threonine (T/2), arginine (R/12) and alanine (A/15) at identical positions in distantly related species such as *A. thaliana* and *L. nivea* suggests that these aminoacids have a conserved function for the loop1 region. To investigate if their substitution affects CENH3 targeting, one of the codifying nucleotides of each of the triplets was exchanged in order to substitute the corresponding aminoacid:

Original:		**Mutated:**	
ACC = Thr, T	polar, neutral, non-aromatic	AAC = Asn, N	polar, neutral, acidic
CGU = Arg, R	polar, positively charged, basic	CCU = Pro, P	non-polar, neutral, aliphatic
GCU = Ala, A	non-polar, neutral,aliphatic	GUU = Val, V	non-polar, neutral, aliphatic

4.2.1 Mutations within the Loop1 region cause diffuse nuclear CENH3 localisation and reduced fertility

Immunostaining with anti-GFP on squashed root tips of AtCENH3(M) transformed plants revealed diffuse signals over the nucleoplasm (Fig. 20 A-F), indicating that the mutated protein lost the ability to target the centromere.

Dispersed signals were visualized during interphase, but no signals were found on metaphase chromosomes (Fig. 20 G-I).

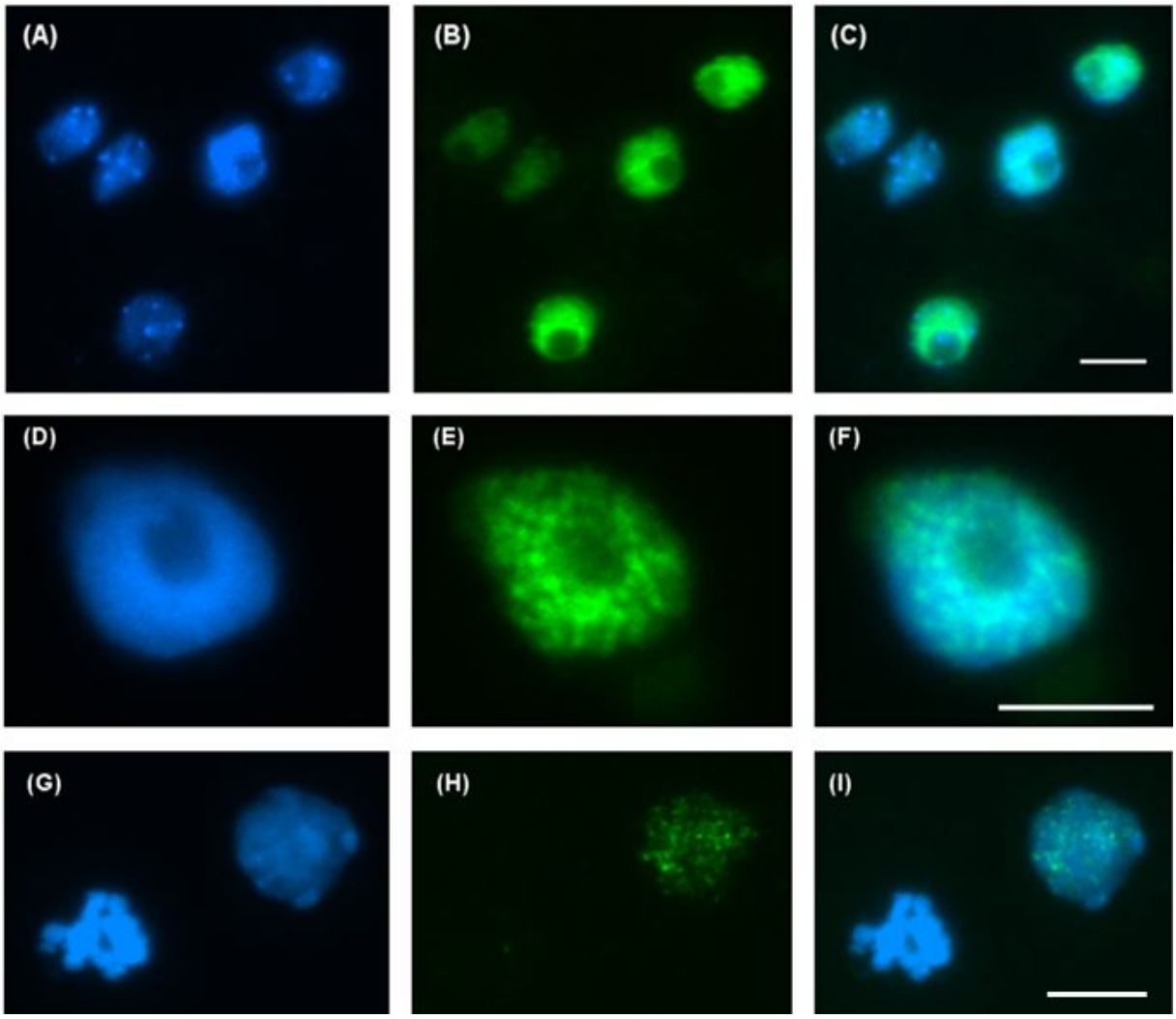

FIGURE 20 Immunostaining on root tip nuclei of lines AtCENH3(M)-5 **(A-C)** and AtCENH3(M)-10 **(D-F)** Diffuse nuclear immunosignals appear after substituting the three conserved amino acids of the loop1 region, indicating loss of centromere targeting ability; **(G-I)** During mitosis, the signals are absent. Blue = DAPI; green = anti-GFP; bar = 5µm.

The expression levels of the recombinant protein and of endogenous CENH3 were compared by Western blotting with anti-GFP and anti-CENH3 antibodies (Fig. 21). For the endogenous CENH3, a band of ~26 kDa was detected, althought the expected size is 18 kDa. Due to their highly basic nature, histones

migrate on gels slower than neutral proteins, what might explain their apparently larger-than-expected molecular weight.

The endogenous CENH3 showed a much higher expression level in comparison to the mutated recombinant variant in the majority of the studied lines.

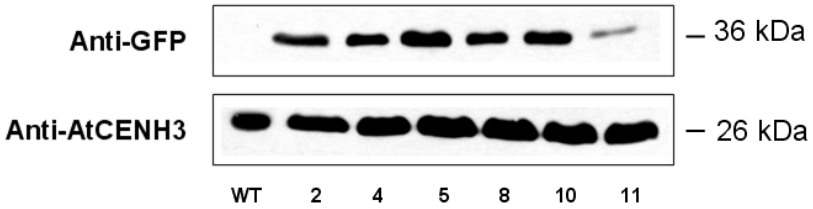

FIGURE 21: Western blotting with anti-GFP confirmed the translation of the recombinant protein. Anti-AtCENH3 revealed stronger expression of the endogenous than of the recombinant CENH3

Despite individual differences as to the expression level, all the selected lines showed the same diffuse immunostaining pattern. The vegetative phenotype of transgenic plants expressing AtCENH3(M) is WT like, but most of them have few and smaller anthers with less viable pollen in comparison to WT. Alexander staining of anthers revealed a high percentage of unviable pollen grains in some of the lines (Fig. 22 B) in comparison to the WT (Fig. 22 A), although lines such as AtCENH3-5 had mostly viable pollen grains. Counting of aborted seeds revealed variation between lines. Line 4 displayed 12% of aborted seeds, line 5, 7%, line 8, 6%, compared to 4% found in WT siliques. Line 10 had with 52% the highest proportion of aborted seeds. The presence of a single stem (Fig. 22 E) was observed in all the lines, even those with low percentage of aborted seeds.

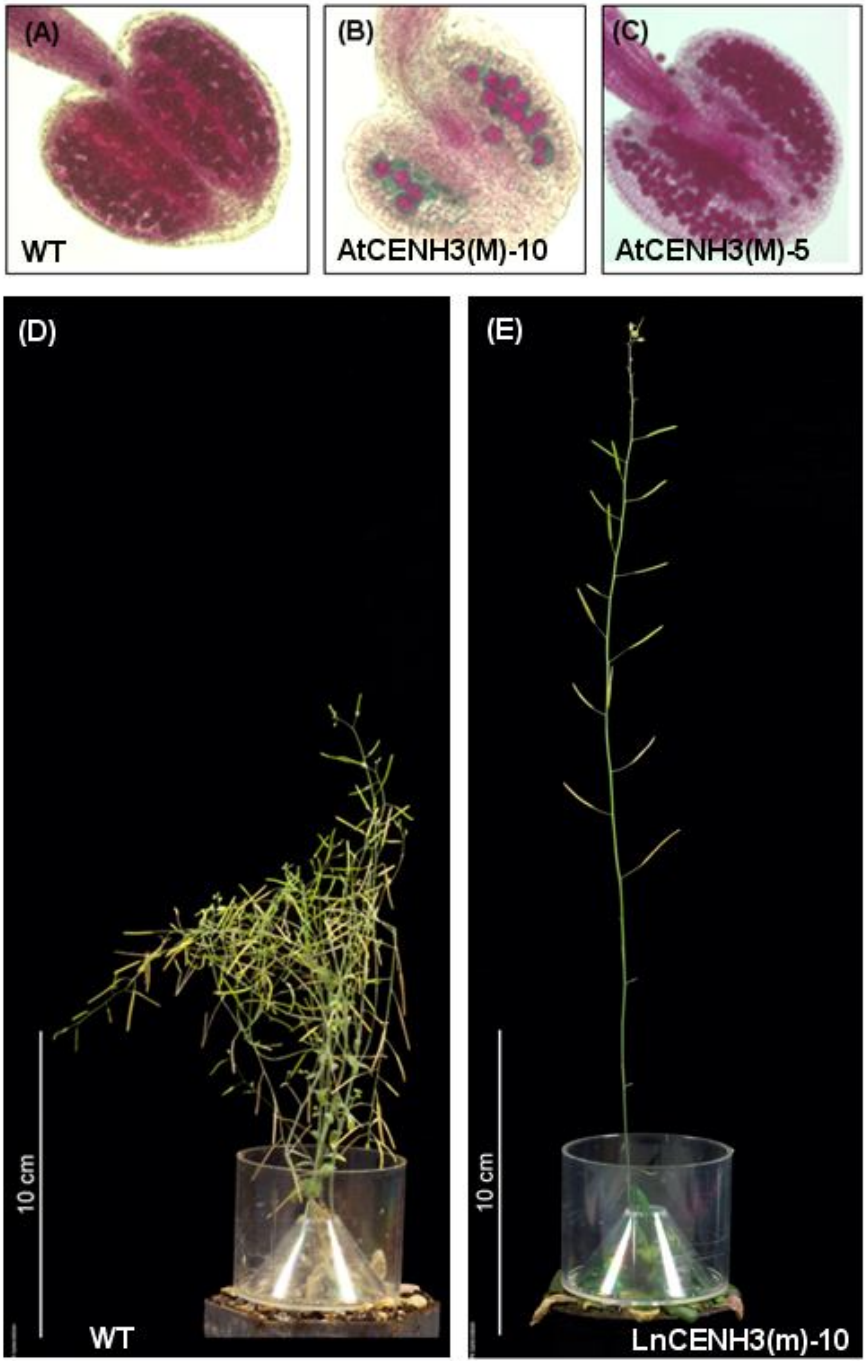

FIGURE 21 AtCENH3(M) lines show reduced fertility. **(A)** anthers of *A. thaliana* WT displayed viable pollen grains, while some of the AtCENH3(M) transformed lines such as -10 revealed a high percentage of unviable pollen grains in green **(B)**; however, some lines such as -5 had viable pollen grains set **(C)**; in contrast to the many stems present on WT plant **(D),** smaller anthers and siliques and only one single stem were common features of AtCENH3(M) lines **(E)**.

5. Discussion

5.1 Transient expression proved expression of recombinant CENH3 but is not reliable for protein expression studies

Barley leaves bombarded with gold particles carrying AarCENH3, AlyCENH3, AtCENH3, CapCENH3, ZmCENH3, LnCENH3-A, LnCENH3-B and HvCENH3 confirmed the expression of all the constructs used in the present work. However, the bombarded cells displayed different localization patterns than that ones found in stably transformed *A. thaliana* plants.

In contrast to stable transformants which constitutively carry the transgene in every cell, constructs transiently expressed by particles bombardment are only able to show fluorescence signals in the cells targeted by the gold particles. Furthermore, recombinant CENH3 would only be loaded if the construct target mitotically active cells. As HvCENH3 deposition occurs mainly in late G2 (table 1), likely most of the bombarded cells were already differentiated, therefore unable to incorporate any CENH3 into centromeric nucleosomes. The same can be hold for the lack of AtCENH3 targeting in barley cells: since AarCENH3 and AlyCENH3 yielded centromeric signals, AtCENH3 should do as well. Thus, although a powerful tool that gives fast results, especially in the case of cultured human and *Drosophila* cells, transient transformation experiments in plant organs reveal expression but not reliably function of recombinant CENH3.

5.2 A high degree of sequence similarity of alien CENH3 to the endogenous protein is required for targeting to *A. thaliana* centromeres

Previously, the C-terminal part of CENH3, including the loop1 region, was shown to be sufficient for centromere targeting in *A. thaliana* (Lermontova et al. 2006). In mammalian cells, canonical histone H3 carrying the CATD is targeted to the centromeres (Black et al. 2004, 2007).

Despite of the relatively low degree of similarity at the nucleotide level of some of the tested species, all recombinant CENH3 proteins, except that of LnCENH3, were targeted to *A. thaliana* centromeres. Even ZmCENH3 localized at *Arabidopsis* centromeres, although *Z. mays* and *A. thaliana* are distantly related and share only about 48% similarity at protein level. In contrast, Nagaki

et al. (2010) reported that CENH3 from another monocot, *Oriza sativa* tagged with GFP did not localize at the centromeres in cultured cells of *A. thaliana*. Of the three amino acids of loop1 of AtCENH3 that are shared by all CENH3 constructs used in the present work, only two of them are conserved in OsCENH3 (Fig 5).

The immunostaining signals of the recombinant proteins of AarCENH3, AlyCENH3, CbpCENH3 and ZmCENH3 on pre-mitotic and mitotic chromosomes indicate a similar loading pattern as described for AtCENH3 (Lermontova et al. 2006). During G2 double signals appear, indicating normal deposition of the alien CENH3 proteins. Stable transmission to the next generation and expression of the transgene was found in T2.

Three different patterns of nuclear signal distribution were found in root tips of transformed plants expressing CbpCENH3 and ZmCENH3. The majority of nuclei displayed centromeric localization (see table 4), but ~24% of the nuclei of both transformants displayed diffuse and centromeric signals, while 15% and 20% respectively were only diffuse. Previously it was shown that AtCENH3 can target *A. lyrata* centromeres but not enter *V. faba* nuclei (Lermontova et al. 2006). This indicates that a high degree of conservation is needed for an efficient CENH3 targeting in *A. thaliana*. AarCENH3 and AlyCENH3 are very closely related to AtCENH3. Possibly AtCENH3 loading factors recognize and deliver these recombinant proteins for correct kinetochore assembly. In contrast, the more distantly related proteins CpbCENH3 and ZmCENH3 could be loaded but with a lower efficiency, at least when AtCENH3 is available. If the endogenous CENH3 is preferentially recruited by chaperones, possibly much of the not deposited alien protein remains in the nucleoplasm until being degraded by proteolysis. This would explain the relatively high percentage of diffusely labeled nuclei in lines expressing these alien proteins. To find out whether the recombinant CENH3 is able to functionally compensate endogenous, AtCENH3 should be silenced within the transformed lines.

5.3 Differences between nuclear importing factors of species with monocentric and species with holocentric chromosomes might result in cytoplasm localization of recombinant LnCENH3 protein

Recombinant EYFP-ZmCENH3 may target *A. thaliana* centromeres, albeit with less efficiency than the more closely related AlyCENH3 and AarCENH3. Despite having the same similarity degree as ZmCENH3 to AtCENH3, LnCENH3 could not be transported into *A. thaliana* nuclei.

De novo chromatin assembly requires deposition of newly synthesized core histones (Mühlhäusser et al. 2001). Bidirectional transport of histones between nucleus and cytoplasm is mediated by transport factors known as importins and exportins (reviewed in Weis et al. 1998). Mühlhäusser et al. (2001) demonstrated that different importins interact directly with individual histone cores and the import efficiency of histones into the nucleus varies according to the availability of importins. H1 for instance could not be transported into the nucleus by any of the importins which transport H2A, H2B, H3 and H4. This suggests a preferential interaction between transporters and histones, which can be the reason of the non-transportation of recombinant LnCENH3 into *A. thaliana* nucleus.

Using *Arabidopsis* and tobacco BY-2 cell cultures, Nagaki et al. (2010) investigated the centromere-targeting ability of GFP-tagged full length CENH3 of *Nicotiana tabacum*, *A. thaliana*, *L. nivea* and *Oriza sativa*. NtCENH3 and AtCENH3 were found at Arabidopsis and tobacco centromeres, and OsCENH3 yielded mainly diffuse signals. Nagaki et al. (2010) found cells transformed with GFP-LnCENH3-A displaying dispersed signals and occasionally centromere-specific ones. Similarly, in the present work, EYFP-LnCENH3-B transiently expressed in barley leaves displayed dispersed signals. Possibly, in stably transformed *A. thaliana* plants, a more tight regulation than in cultured cells suppresses entering of nuclei and deposition at centromeres of the less related LnCENH3.

5.4 The three mutated aminoacids of AtCENH3 are crucial but not sufficient for centromere targeting

The loop1 region was shown to be necessary and sufficient for CENH3 targeting, but it remained to be established what interactions are responsible for loop1 function. Loop1 makes contacts with the DNA. After substituting the three amino acids which the tested species share within their loop 1 region, T/2, R/12 and A/15 by N/2, P/12 and V/15, the mutated AtCENH3 was no longer able to target *A. thaliana* centromeres. Apparently, the arginine or even all these three conserved amino acids of the loop 1 region are needed for interaction with DNA, with chaperones or other proteins. Veermak et al. (2002) found the conservation of the loop1 region to be critical for centromere targeting in *Drosophila.*

Although the mutated AtCENH3 can not be loaded, it enters *A. thaliana* nucleus, while LnCENH3 localized only at the cytoplasm. This means, the substitution of the three conserved amino acids disturbs centromere targeting, but their presence alone (as in LnCENH3) does not guarantee correct loading. Therefore, the tested 3 amino acids of the loop1 region are necessary but not sufficient for CENH3 incorporation into *A. thaliana* centromeric nucleosomes *in planta,* and for entering the nucleus apparently other amino acid combinations are required.

5.5 Mutated AtCENH3 impairs fertility

Although the transformed plants showed no strongly deviating vegetative phenotype, most of the lines carrying the mutated sequence were at least partially sterile. Alexander staining of pollen grains revealed a much higher number of unviable pollen grains in transgenic than in WT plants. Plants with many unviable pollen grains have also a high percentage of seeds abortion. Recently Lermontova et al. (submitted) demonstrated that transgenic plants expressing a CENH3-RNAi construct exhibit meiotic defects and reduced fertility indicating that CENH3 is also essential for correct meiotic divisions.

One possible reason for the impaired fertility is that the expression of a recombinant EYFP-AtCENH3(M) decreases the expression of the endogenous CENH3 during meiosis *via* meiosis-specific RNA interference. Alternatively, the

substitution of T/2, R/12 and A/15, localized at the CATD, which has role in preventing its mislocalization to euchromatin via degradation (Ranijtkar et al. 2010), could prevent protein degradation if the original aminoacids are important for targeting of proteolysis machinery.

At present it is not clear whether CENH3 has an additional role during meiosis, but many observations point into this direction. CENH3-RNAi lines of *A. thaliana* and of transgenic lines EYFP-AtCENH3 with only the C-terminal part of AtCENH3 (Lermontova et al. submitted) had disturbances in pollen grain production, ovule development and consequently a high number of aborted seeds. Reduced fertility of these lines further support the assumption of a not yet defined meiotic function of CENH3.

5.6 *L. nivea* CENH3 genes are transcribed; the functionality of LnCENH3-A remains to be proven

In genomes carrying a gene duplication, often one of them gets silenced, but eventually both duplicated genes are expressed. In two allotetraploid *Oriza* species, two CENH3 genes were identified, and both are transcribed, showing no preferential expression of one of them (Hirsch et al. 2009).

However, also for diploid species, such as *A.halleri* and *A. lyrata,* two CENH3 genes were reported (Kawabe et al. 2006). Both are expressed, but is not clear whether they are also functional. Multiple CENH3 genes were found also in metazoan. The nematode *Caenorhabditis elegans* has two highly homologous CENH3 loci, Hcp-3 and Cpar-1, but the latter one is weakly expressed (Monen et al. 2005).

LnCENH3-A and B were isolated from cDNA of the *bona fide* diploid *L. nivea*, what means both isoforms are transcribed. The incorporation into centromeric nucleosomes is clear only for the B isoform. The peptide used to raise anti-LnCENH3-A is shared by both isoforms. Since anti-LnCENH3-A may cross recognize LnCENH3-B, it is not yet clear whether LnCENH3-A is active. Additional experiments with isoforms specific antibodies, are necessary to investigate co-localization of both proteins.

6. Conclusions

The results obtained in this work led to the main conclusions:

- A high similarity degree to the endogenous protein is required for an efficient targeting of alien CENH3 to *A. thaliana* centromeres;

- The threonine (T/2), arginine (R/12) and alanine (A/15) residues at loop1 region of CENH3 are necessary but not sufficient for targeting the centromeres;

- *Luzula nivea* has two CENH3 isoforms (A and B) and both might be functional.

7. Outlook

1. To investigate whether alien CENH3 can functionally substitute AtCENH3, the following experiments should be done:
    - Silencing of the endogenous AtCENH3 of *Arabidopsis thaliana* lines carrying *Alien*CENH3 to figure out if the recombinant protein is able to rescue the plants;

2. To test if one aminoacid or a combination of the three is required for CENH3 targeting to the centromeres and to understand how the substitution of the conserved aminoacids of the loop 1 region impairs fertility:
    - *Arabidopsis* WT shall be transformed with CENH3 constructs carrying single mutated amino acids at the loop1 region;
    - Plants with different sterility degrees must be analyzed for presence, localization and relative amount of the recombinant protein during meiosis;

3. To characterize LnCENH3-B and to investigate differences in the distribution of –A and –B isoforms during mitotic and/or meiotic division, the following work should be carried:
    - Generation of anti-LnCENH3-A and application for immunostaining and western blotting experiments.

8. Summary

Eukaryotic chromosomes need centromeres to ensure their faithful transmission to daughter nuclei during nuclear division. The centromere is the site where the kinetochore assembles for chromosome attachment to the spindle microtubules, directing the chromosome segregation during nuclear division. Kinetochore assembly requires deposition of the centromeric histone H3 variant (CENH3) into centromeric nucleosomes. CENH3 has a variable N-terminal and a more conserved C-terminal part, including the loop1 region of the histone fold domain, which is considered to be critical for centromere targeting. To investigate the structural requirements for centromere targeting, constructs for tagged CENH3 of *A. lyrata, A. arenosa*, *Capsella bursa-pastoris*, *Zea mays* and *Luzula nivea* (the latter with holocentric chromosomes) were transformed into *A. thaliana*. Except for LnCENH3, all recombinant CENH3 proteins targeted *A. thaliana* centromeres, but the more distantly related the alien protein is, the lower is the efficiency of targeting. Alignment of CENH3 sequences revealed that the tested species share only three amino acids at loop1 region: threonine (T/2), arginine (R/12) and alanine (A/15). These three amino acids were substituted by asparagine, proline and valine encoding sequences within a recombinant EYFP-AtCENH3 construct via PCR mutagenesis prior to transformation of *A. thaliana*. After transformation, immunostaining of root tip nuclei with anti-GFP antibodies yielded only diffuse signals, indicating that the original three amino acids are necessary but not sufficient for targeting *A. thaliana* centromeres.

9. Zusammenfassung

Eukaryotische Chromosomen benötigen ein Zentromer zu ihrer korrekten Verteilung auf die Tochterkerne während der Kernteilung. Am Zentromer wird das Kinetochore gebildet, das die Spindefaseransatzstelle darstellt und die Verteilung der Chromosomen während der Kernteilung steuert. Die Deposition der centromerischen Histonvariante CENH3 in die zentromerischen Nukleosomen ist eine Voraussetzung für die Kinetochorbildung. CENH3 besteht aus einem variablen N-terminalen und einem mehr konservierten C-terminalen Teil. Der C-terminale Teil schließt die Loop 1-Region der Histonfaltungsdomäne ein, die für die CENH3-Deposition am Zentromer erforderlich ist. Zur Untersuchung der strukturellen Voraussetzungen für die Zentromerbeladung mit CENH3 wurden Konstrukte für fluoreszenz-markierte CENH3-Proteine aus *A. lyrata, A. arenosa, Capsella bursa-pastoris, Zea mays* und *Luzula nivea* (letztere mit holozentrischen Chromosomen) erstellt und damit *A. thaliana* transformiert. Außer Lu LnCENH3 gelangten alle rekombinanten CENH3-Proteine in die Zentromeren von *A. thaliana*, jedoch nahm die Effizienz der Zentromerbeladung mit dem phylogenetischen Abstand ab.

Ein Vergleich der Proteine der getesteten Arten ergab eine Übereinstimmung von nur 3 Aminosäureresten zwischen den entsprechenden Loop 1-Regionen: Threonin (T2), Arginin (R12) und Alanin (A15). Diese 3 wurden gemeinsam mittels PCR-Mutagenese durch Asparagin, Prolin und Valin kodierende Sequenzen in einem rekombinanten EYFP-AtCENH3 Konstrukt ersetzt.

A. thaliana-Pflanzen, die mit diesem Konstrukt transformiert waren ergaben nach Immunfärbung von Kernen aus Wurzelspitzen mit anti-GFP-Antikörpern lediglich diffuse Fluoreszenz-Signale.

Offenbar sind die substituierten 3 Aminosäuren nötig aber nicht ausreichend (s. Lu CENH3) um die zentromerischen Nukleosomen von *A. thaliana* mit CENH3 zu beladen.

10. References

Ahmad K., Henikoff S. (2001) Modulation of a transcription factor counteracts heterochromatic gene-silencing in Drosophila. *Cell* 104, 839-847

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. (2002) Molecular Biology of the Cell. 4th edition. New York: Garland Science

Alexander M. P. (1969) Differential staining of aborted and nonaborted pollen. *Stain Technol*. 44, 117-122

Amor D.J., Kalitsis P., Sumer H., Choo K.H. (2004). Building the centromere: From foundation proteins to 3D organization. *Trends Cell Biol*. 14, 359–368

Aragon-Alcaide L., Miller T., Schwarzacher T., Reader S., Moore G. (1996) A cereal centromeric sequence. *Chromosoma* 105, 261–268

Baskin T.I., Cande W.Z**.** (1990) The structure and function of the mitotic spindle in flowering plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol*. 41, 277-315

Baum M., Ngan V., Clarke L. (1994) The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol. Biol. Cell* 5, 747–761

Bechtold N., Ellis J., Pelletier G. (1993). In planta *Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Life Sci*. 316, 1194–1199.

Bertani, G. (1951) Studies on lysogenesis . I . The mode of phage liberation by lysogenic *Escherichia coli* . *J . Bacteriol .* 62, 293-300

Biggins S., Severin F. F., Bhalla N., Sassoon I., Hyman A. A., Murray A. W. (1999) The conserved protein kinase Ipl1 regulates microtubule binding to kinetochores in budding yeast. *Genes Dev*. 13(5), 532-44

Bischoff J. R., Anderson L., Zhu Y., Mossie K., Ng L., Souza B., Schryver B., Flanagan P., Clairvoyant F., Ginther C., Chan C. S., Novotny M., Slamon D. J., Plowman G. D. (1998) A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J.* 17, 3052-3065

Black B. E., Foltz D. R., Chakravarthy S., Luger K., Woods V. L., (2004) Structural determinants for generating centromeric chromatin. *Nature* 430, 578–582

Black B. E., Brock M. A., Bedard S., Woods V. L. Jr, Cleveland D. W. (2007) An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. *Proc Natl Acad Sci USA* 104, 5008–5013

Blower M. D., Sullivan B. A., Karpen G. H. (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* 2, 319–330

Blower M. D., Daigle T., Kaufman T., Karpen G. H. (2006) *Drosophila* CENP-A mutations cause a BubR1-dependent early mitotic delay without normal localization of kinetochore components. *PLoS Genet* 2, e110

Bradford M. (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem*. 72, 248-254

Buchwitz B.J., Ahmad K., Moore L.L., Roth M.B., Henikoff S. (1999) A histone-H3-like protein in *C. elegans*. *Nature.* 401, 547–548

Caillaud M. C., Paganelli L., Lecomte P., Deslandes L., Quentin M., Pecrix Y., Le Bris M., Marfaing N., Abad P., Favery B. (2009) Spindle Assembly

Checkpoint Protein Dynamics Reveal Conserved and Unsuspected Roles in Plant Cell Division. *PLoS* 4(8), e6757

Chan G. K., Liu S-T. Yen T. J. (2005) Kinetochore structure and function. *Trends in Cell Biology* 15(11), 589-598

Cheeseman I. M., Desai A. (2008) Molecular architecture of the kinetochore–microtubule interface. *Nature Reviews Molecular Cell Biology* 9, 33-46

Chen Y., Baker R. E., Keith K. C., Harris K., Stoler S. (2000) The N terminus of the centromere H3-like protein Cse4p performs an essential function distinct from that of the histone fold domain. *Mol. Cell. Biol*. 20, 7037–7048

Choo, K. H. A. (1997). The Centromere. Oxford, UK: *Oxford University Press*.

Choo K. H. A. (2001) Domain organization at the centromere and neocentromere. *Dev. Cell* 1, 165–177

Clarke, L. (1990) Centromeres of budding and fission yeasts. *Trends Genet*. 6, 150-154

Collins K.A., Furuyama S., Biggins S. (2004) Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. *Curr. Biol.* 14, 1968–1972

Cooper J. L., Henikoff S. (2004) Adaptative evolution of the histone fold domain in centromeric histones. *Mol. Biol. Evol.* 21, 1712-1718

Dalal Y., Bui M. (2010) Down the rabbit hole of centromere assembly and dynamics. *Current Opinion in Cell Biology* 22, 1–11

Dawe R. K., Reed L., Yu H. G., Muszynski M. G., Hiatt E. N. (1999) A maize homolog of mammalian CENPC is a constitutive component of the inner kinetochore. *Plant Cell* 11, 1227–1238

Dawe R. K., Richardson E.A., Zhang X. (2005) The simple ultrastructure of the maize kinetochore fits a two-domain model. *Cytogenet. Genome Res.* 109, 128–133

Demidov D., Hesse S., Tewes A., Rutten T., Fuchs J., Ashtiyani R. K., Lein S., Fischer A., Reuter G., Houben A. (2009) Aurora1 phosphorylation activity on histone H3 and its cross-talk with other post-translational histone modifications in Arabidopsis. *Plant J.* 59, 221-230

Dernburg, A.F. (2001). Here, there, and everywhere: Kinetochore function on holocentric chromosomes. *J. Cell Biol.* 153, F33–F38

Dimitriadis E. K., Weber C., Gill R. K., Diekmann S., Dalal Y. (2010) Tetrameric organization of vertebrate centromeric nucleosomes. *Proc. Natl. Acad. Sci. USA* 107, 20317-20322

Du Y., Dawe R. K. (2007) Maize NDC80 is a constitutive feature of the central kinetochore. *Chromosome Res* 15, 767–775

Dubin M, Fuchs J, Gra¨f R, Schubert I, Nellen W (2010) Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis. *Nucleic Acids Res.* (in press)

Dunleavy E. M., Roche D., Tagami H., Lacoste N., Ray-Gallet D., Nakamura Y., Daigo Y., Nakatani Y., Almouzni-Pettinotti G. (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* 137, 485–497

Earnshaw W. C., and Rothfield N. (1985) Identification of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 91, 313–321

Edwards K. , Johnstone C. , Thompson C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucl Acid Res* 19, 1349

English C. M., Adkins M. W., Carson J. J., Churchill M. E., Tyler J. K. (2006) Structural basis for the histone chaperone activity of Asf1. *Cell* 127(3), 495-508

Erhardt S., Mellone B. G., Betts C. M., Zhang W., Karpen G. H., et al. (2008) Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. *Journal of Cell Biology* 183, 805–818

Foltz D. R., Jansen L. E., Bailey A. O., Yates Y. R., Bassett E. A., Wood S., Black B. E., Clevel D. W. (2009) Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell 137, 472-484.*

Francisco L., Wang W., Chan C. S. (1994) Type 1 protein phosphatise acts in opposition to IpL1 protein kinase in regulating yeast chromosome segregation. *Mol. Cell. Biol.* 14, 4731-4740

Fujita Y., Hayashi T., Kiyomitsu T., Toyoda Y., Kokubu A., Obuse C., Yanagida M. (2007) Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev. Cell* 12, 17–30

Furuyama T., Dalal Y., Henikoff S. (2006) Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc. Natl. Acad. Sci. USA.* 103, 6172–6177

Gascoigne K. E., Cheeseman I. M. (2010) Kinetochore assembly: if you build it, they will come. Curr Opin Cell Biol. 23, 1-7

Gill N., Hans C., Jackson S. A. (2008) An overview of plant chromosome structure. *Cytogenetic and Genome Res.* 120, 194-201

Glover D. M., Leibowitz M. H., McLean D. A., Parry H. (1995) Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. *Cell* 81, 95-105

Hall S. E., Kettler G., Preuss D. (2003) Centromere satellites from Arabidopsis populations: Maintenance of conserved and variable domains. *Genome Res* 13, 195–205

Hanahan D. (1983) Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* 166(4)**,** 557-580

Hayashi T., Fujita Y., Iwasaki O., Adachi Y., Takahashi K., Yanagida M. (2004) Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres, *Cell* 118, 715–729.

Heeger S., Leismann O., Schittenhelm R., Schraidt O. Heidmann S., Lehner C. F. (2005) Genetic interactions of separase regulatory subunits reveal the diverged Drosophila Cenp-C homolog. *Genes Dev* 19, 2041–2053

Hegemann J. H., Shero J. H., Cottarel G., Philippsen P., Hieter P. (1988) Mutational analysis of centromere DNA from chromosome VI of *Saccharomyces cerevisiae. Mol Cell Biol* 8, 2523–2535

Hein, J. (1990) Reconstructing evolution of sequences subject to recombination using parsimony. *Math. Biosci*. 98, 185-200

Hemmerich P., Weidtkamp-Peters S., Hoischen C., Schmiedeberg L., Erliandri I., Diekmann S. (2008) Dynamics of inner kinetochore assembly and maintenance in living cells. *J. Cell Biol.* 180, 1101–1114

Heng H. H., Krawetz S. A., Lu W., Bremer S., Liu G., Ye C. J. (2001) Re-defining the chromatin loop domain. *Cytogenet. Cell Genet.* 93, 155-161

Hennig L., Taranto P., Walser M., Schönrock N., Gruissem W. (2003) *Arabidopsis* MSI1 is required for epigenetic maintenance of reproductive development. *Development* 130, 2555–2565

Hirsch C. D., Wu Y., Yan H., Jiang J. (2009) Lineage-specific adaptive evolution of the centromeric protein CENH3 in diploid and allotetraploid *Oryza* species. *Mol Biol Evol* 26, 2877–2885

Houben A., Schubert I. (2003) DNA and proteins of plant centromeres. *Curr. Opin. Plant Biol.* 6, 554–560

Howman E. V., Fowler K. J., Newson A. J., Redward S., MacDonald A. C., Kalitsis P., Choo, K. H. A. (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc. Natl. Acad. Sci. USA* 97, 1148–1153

Hoyt M. A., Totis L., Roberts B. T. (1991) *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell* 66, 507–517

Ishii K. (2009) Conservation and divergence of centromere specification in yeast. *Curr. Opin. Microbiol.* 12(6), 616–622

Jansen L. E., Black B. E., Foltz D. R., Cleveland D. W. (2007) Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* 176, 795-805

Jasencakova Z., Meister A., Walter J., Turner B.M., Schubert I. (2000). Histone H4 acetylation of euchromatin and heterochromatin is cell cycle dependent and correlated with replication rather than with transcription. *Plant Cell* 12, 2087–2100

Jehn B., Niedenthal R., Hegemann J. H. (1991) In Vivo Analysis of the Saccharomyces cerevisiae Centromere CDEIII Sequence: Requirements for Mitotic Chromosome Segregation. *Mol Cell Biol*, 5212-5221

Jiang J., Nasuda S., Dong F., Scherrer C.W., Woo S.-S., Wing R.A., Gill B.S., Ward D.C. (1996) A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *Proc. Natl. Acad. Sci. USA* 93, 14210–14213

Kawabe A., Nasuda S., Charlesworth D. (2006) Duplication of Centromeric Histone H3 (HTR12) Gene in Arabidopsis halleri and A. lyrata Plant Species With Multiple Centromeric Satellite Sequences. *Genetics* 174(4), 2021 - 2032.

Kaya H., Shibahara K. I., Taoka K. I., Iwabuchi M., Stillman B., Araki T. (2001) *FASCIATA* genes for chromatin assembly factor-1 in *Arabidopsis* maintain the cellular organization of apical meristems. *Cell* 104, 131–142

Kimbara J., Endo T. R., Nasuda S. (2004) Characterization of the genes encoding for MAD2 homologues in wheat. *Chromosome Res* 12, 703–714

Koncz C., Schell J. (1986) The promoter of the T-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383–396

Kwon M. S., Hori T., Okada M., Fukagawa T. (2007). CENP-C Is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. *Mol. Biol. Cell* 18

Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 (5259), 680-685

Lam A. L., Pazin D. E., Sullivan B. A. (2005) Control of gene expression and assembly of chromosomal subdomains by chromatin regulators with antagonistic functions. *Chromosoma* 114, 242-251

Lamb J. C., Yu W., Han F., Birchler J. A. (2007) Plant chromosomes from end to end: Telomeres, heterochromatin and centromeres. *Curr. Opin. Plant Biol*. 10**,** 116–122

Lampson M. A, Cheeseman I. M. (2010) Sensing centromere tension: Aurora B and the regulation of kinetochore function. *Trends Cell Biol*., in press

Lermontova I., Schubert V., Fuchs J., Klatte S., Macas J., Schubert I. (2006) Loading of *Arabidopsis* centromeric histone CENH3 occurs mainly during G2 and requires the presence of the histone fold domain. *Plant Cell* 18, 2443–2451

Lermontova I., Fuchs J., Schubert V., Schubert I. (2007) Loading time of the centromeric histone H3 variant differs between plants and animals. *Chromosoma* 116, 507–510

Lermontova I., Fuchs J., Schubert I. (2008) The Arabidopsis checkpoint protein Bub3.1 is essential for gametophyte development. *Front Biosci* 13, 5202–5211

Li R., Murray A. W. (1991) Feedback control of mitosis in budding yeast. *Cell* 66, 519–531

Maddox P. S., Oegema K., Desai A., and Cheeseman I. M. (2004). "Holo"er than thou: Chromosome segregation and kin**et**ochore function in *C. elegans*. *Chromosome Res.* 12, 641–653

Malik H. S., Henikoff S. (2001) Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. *Genetics* 157, 1293–1298

Malik H. S., Henikoff S. (2002) Conflict begets complexity: the evolution of centromeres. *Curr. Opin. Genet. Dev.* 12, 711–718

Martinez-Zapater J., Estelle M. A., Somerville C. R. (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol Gen Genet* 204, 417–423

Maruyama S., Kuroiwa H., Miyagishima S. Y., Tanaka K., Kuroiwa T. (2007) Centromere dynamics in the primitive red alga *Cyanidioschyzon merolae*. *Plant J.* 49, 1122–1129

May B. P., Lippman Z. B., Fang Y., Spector D. L., Martienssen R. A. (2005) Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. *PLoS Genet.* 1 (6), e79

McCombie W. R, de la Bastide M., Habermann K., Parnell L. D., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Yordan C. (2000) The complete sequence of a heterochromatic island from a higher eukaryote. *Cell* 100**,** 377-386

McGrew J., Diehl, B., Fitzgerald-Hayes M. (1986) Single base-pair mutations in centromere element III cause aberrant chromosome segregation in *Saccharomyces cerevisiae. Mol Cell Biol* 6, 530–538

Meluh P. B., Yang P., Glowczewski L., Koshland D., Smith M. M. (1998) Cse4p is a component of the core centromere of *Saccharomyces cerevisiae Cell* 94, 607-613

Miller J. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Monen J., Maddox P.S., Hyndman F., Oegema K., Desai A. (2005) Differential role of CENP-A in the segregation of holocentric *C. elegans* chromosomes during meiosis and mitosis. *Nat. Cell Biol.* 12, 1248–1255

Moore L.L., Morrison M., Roth M.B. (1999) HCP-1, a protein involved in chromosome segregation, is localized to the centromere of mitotic chromosomes in *Caenorhabditis elegans*. *J. Cell Biol.*147, 471–480

Moreno-Moreno O., Torras-Llort M., Azorín F. (2006) Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of *Drosophila*, to centromeres. *Nucleic Acids Res.* 34, 6247–6255

Moroi Y., Peebles C., Fritzler M. J., Steigerwald J. Tan E. M (1980) Autoantibody to centromere (kinetochore) in scleroderma sera. *Proc Natl Acad Sci USA* 77, 627–1631

Morris C. A., Moazed D. (2007) Centromere assembly and propagation. *Cell* 128(4), 647-50

Mühlhäusser P., Muller E. C., Otto A., Kutay U. (2001) Multiple pathways contribute to nuclear import of core histones. *EMBO Rep.* 2, 690–696.

Murashige T., Skoog F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant* 15, 473–497

Murray A. W. (1992) Creative blocks: cell-cycle checkpoints and feedback controls. *Nature* 359, 599-604

Musacchio A., Salmon E. D. (2007) The spindle-assembly checkpoint in space and time. *Nat. Rev. Mol. Cell Biol.* 8, 379–393

Nagaki K., Kashihara K., Murata M. (2005) Visualization of diffuse centromeres with centromere-specific histone H3 in the holocentric plant *Luzula nivea*. *Plant Cell* 17,1886-1893

Nagaki K., Terada K., Wakimoto M., Kashihara K., Murata M. (2010) Centromere targeting of alien CENH3s in Arabidopsis and tobacco cells. *Chromosome Res.* 18(2), 203-11

Nicklas R. B., Koch C. A**.** (1969) Chromosome micromanipulation III. Spindle fiber tension and the reorientation of mal-oriented chromosomes. *J. Cell Biol.* 43, 40-50

Nicklas R. B., Ward S. C. (1994) Elements of error correction in mitosis: microtubule capture, release, and tension*. J. Cell Biol.* 126, 1241–1253

Ogura Y., Shibata F., Sato H., Murata M. (2004) Characterization of a CENP-C homolog in *Arabidopsis thaliana. Genes Genet. Syst.* 79,139–144

Palmer D. K., O'Day K., Trong H. L., Charbonneau H., Margolis R. L. (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci. USA* 88*,* 3734–3738

Philpott A., Krude T., Laskey R.A. (2000) Nuclear chaperones*. Semin. Cell. Dev. Biol.* 11, 7–14

Pidoux A. L., Choi E. S., Abbott J. K., Liu X., Kagansky A., Castillo A. G., Hamilton G. L., Richardson W., Rappsilber J., He X. (2009) Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. *Molecular cell* 33,299–311

Presting G. G., Malysheva L., Fuchs J., Schubert, I. (1998) A Ty3/gypsy retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. *Plant J*. 16 , 721–728

Ranjitkar P., Press M. O., Yi X., Baker R., MacCoss M. J., Biggins S. (2010) An E3 Ubiquitin Ligase Prevents Ectopic Localization of the Centromeric Histone

H3 Variant via the Centromere Targeting Domain. *Molecular Cell* 40(3), 455-464

Ridgway P., Almouzni G. (2000) CAF-1 and the inheritance of chromatin states: at the crossroads of DNA replication and repair. *J. Cell Sci.* 113, 2647–2658

Ris H., Witt P. L. (1981) Structure of the Mammalian Kinetochore. *Chromosoma* 82, 153-170

Santaguida S., Musacchio A. (2009) The life and miracles of kinetochores. *EMBO J.* 28, 2511-31

Sato H., Shibata F., Murata M. (2005) Characterization of a Mis12 homologue in *Arabidopsis thaliana*. *Chromosome Res.* 13, 827–834.

Schaar B. T., Chan G. K., Maddox P., Salmon E. D., Yen T. J. (1997) CENP-E function at kinetochores is essential for chromosome alignment. *J. Cell Biol.* 139, 1373–1382

Schueler M., Higgins A., Rudd M., Gustashaw K., Willard H. (2001) Genomic and genetic definition of a functional human centromere. *Science* 294, 109–114

Schuh M., Lehner C. F., Heidmann S. (2007) Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.* 17, 237-243

Schumacher J. M., Ashcroft N., Donovan P. J., Golden A. (1998) A highly conserved centrosomal kinase, AIR-1, is required for accurate cell cycle progression and segregation of developmental factors in Caenorhabditis elegans embryos. *Development* 125, 4391-4402

Sears D. D., Hegemann J. H., Shero J. H., Hieter P. (1995) Cis-acting determinants affecting centromere function, sister-chromatid cohesion and

reciprocal recombination during meiosis in *Saccharomyces cerevisiae. Genetics* 139, 1159–1173

Shelby R. D., Monier K., Sullivan K. F. (2000) Chromatin assembly at kinetochores is uncoupled from DNA replication. *J. Cell Biol.* 151, 1113–1118

Schenk P. M., Elliott A. R., Manners J. M. (1998) Assessment of transient gene expression in plant tissues using the green fluorescent protein. *Plant Mol. Biol. Rep.* 16, 313–322

Smith S., Stillman B. (1989) Purification and characterization of CAF-1, a human cell factor required for chromatin assembly during DNA replication *in vivo*. *Cell* 58, 15–25

Shibahara K., Stillman B. (1999) Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96, 575–585

Silva M. C., Jansen L. E. (2009) At the right place at the right time: novel CENP-A binding proteins shed light on centromere assembly. *Chromosoma* 118(5), 567-574

Singh R. K., Kabbaj M. H., Paik J., Gunjan A. (2009) Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. *Nat. Cell Biol.* 11(8), 925-933.

Smith S., Stillman B. (1991) Stepwise assembly of chromatin during DNA replication *in vitro*. *EMBO J.* 10, 971–980

Sudakin V., Chan G. K, Yen T. J. (2001) Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2 *J. Cell Biol.* 154, 925–936

Sun X., Wahlstrom J., Karpen G. (1997) Molecular structure of a functional Drosophila centromere. *Cell* 91, 1007–1019

Tagami H., Ray-Gallet D., Almouzni G., Nakatani Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51–61

Takahashi K., Takayama Y., Masuda F., Kobayashi Y., Saitoh S. (2005) Two distinct pathways responsible for the loading of CENP-A to centromeres in the fission yeast cell cycle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360**,** 595–606

Talbert P. B., Masuelli R., Tyagi A. P., Comai L., Henikoff S. (2002) Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. *Plant Cell* 14: 1053–1066

Talbert P. B., Bryson T. D., Henikoff S. (2004) Adaptive evolution of centromere proteins in plants and animals. *J. Biol.* 3, 18

ten Hoopen R., Schleker T., Manteuffel R., Schubert I. (2002) Transient CENP-E like kinetochore proteins in plants. *Chromosome Res.* 10, 561-570

Thompson H. L., Schmidt R., Dean C. (1996) Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acids Res.* 24, 3017–3022

Tomkiel J., Cooke C. A., Saitoh H. Bernat R. L., Earnshaw W. (1994) CENP-C is required for maimaining proper kinetochore size and for a timely transition to anaphase. *The Journal of Cell Biology* 125(3), 531-545

Vermaak D., Hayden H.S., Henikoff S. (2002). Centromere targeting element within the histone fold domain of Cid. *Mol. Cell. Biol*. 22, 7553–7561.

Yu H. G., Muszynski M. G., Kelly Dawe R. (1999) The maize homologue of the cell cycle checkpoint protein MAD2 reveals kinetochore substructure and contrasting mitotic and meiotic localization patterns. *J. Cell Biol.* 145, 425–435

Yu H. G., Hiatt E. N., Dawe R. K. (2000) The plant kinetochore. *Trends Plant Sci.* 5, 543–547

Weis, K. (1998) Importins and exportins: how to get in and out of the nucleus. *Trends Biochem Sci*. 23(5), 185-189

Zeitlin S. G., Patel S., Kavli B., Slupphaug G. (2005) Xenopus CENP-A assembly into chromatin requires base excision repair proteins. *DNA Repair* 4, 760-772

Zeitlin S. G., Baker N. M., Chapados B. R. (2009) Double-strand DNA breaks recruit the centromeric histone CENP-A. *Proc. Natl. Acad. Sci. USA* 106 (37), 15762-15767

Zhong, C. X., Marshall, J. B., Topp, C., Mroczek, R., Kato, A., Nagaki, K., Birchler, J. A., Jiang, J., and Dawe, R.K. (2002) Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. *Plant Cell* 14**,** 2825–283

Zweidler A. (1984) Core histone variants of the mouse: primary structure and differential expression. In: Organization Structure, G.S Regulation, J.L Stein, Stein and W.F Marzluff, Editors, *Histone Genes*, John Wiley & Sons, New York

Publication connected with the thesis

Moraes I. C. R., Lermontova, I, Schubert I. (2011) Recognition of *A. thaliana* centromeres by heterologous CENH3 requires high similarity to the endogenous protein. Plant Molecular Biology, in press.

Curriculum Vitae

PERSONAL IDENTIFICATION

- Name: Izabel Cristina Ribeira de Moraes
- Date of Birth: February, 25, 1979 in São Paulo, Brazil
- Marital Status: Single
- Contact information: Karyotype Evolution group
  Leibniz Institute of Plant Genetics and Crop Plant Research
  Corrensstrasse 3 - 06466 - Gatersleben, Germany
  phone: 0049 (0) 39482 5537
  e-mail: izabel_moraes@yahoo.com

ACADEMIC QUALIFICATIONS

Degree: Master of Science in Genetics, Plant Breeding and Biotecnology
Institution: Agronomy Institute in Campinas
Dissertation title: "Cytogenetic characterization and reproductive biology of three species of genus *Hypericum* L."
Concluded in January/2007

Degree: Bachelor in Biology
College: FACIS – Faculdade de Ciências da Saúde de São Paulo
(Faculty of Health Sciences of São Paulo)
Concluded in March/2004

PROFESSIONAL EXPERIENCE

- Since January 2008 – Leibniz Institute of Plant Genetics and Crop Plant Research
  In my PhD work, I studied some of the requirements for targeting of the centromeric histone H3 to *Arabidopsis thaliana* centromeres. In addition to cytogenetics techniques (squashs, immunostaining, Fluorescence *In Situ* Hibridization, chromatin extended fibers) my dissertation work has involved molecular biology techniques (PCR, cloning, construct preparation, site direct mutagenesis), plant stable and transient transformation and selection, tools for protein and gene analysis such as western and southern blotting and semi quantitative RT-PCR, and data processing with Adobe Photoshop and Acrobat.

- 2005-2007 - Agronomy Institute in Campinas
  During my Master´s I have studied the cytogenetics and reproductive biology of *Hypericum brasiliense*, *H. cordatum* and *H. ternum,* species with medicinal properties and potential agronomic interest.

- 2003-2007 - São Paulo State Government – Department of Education -
  I worked in public schools teaching Sciences and Biology to elementary and secondary schools.

- 2002-2003 - Faculdade de Ciências da Saúde de São Paulo (Faculty of Health Sciences of São Paulo)
  I worked at the Department of Botany performing taxonomical identifications on Herbarium and preparing material for herborization. I also worked under the supervision of Dr. Melânia Cornélio on the composition of essential oils from pomegranate (*Punica granatum*) fruit.

CONGRESSES, MEETINGS AND CONFERENCES

MORAES, I.; LERMONTOVA, I.; SCHUBERT, I. (2010) How is CENH3 recognizing centromeric DNA? Conference: Replication, Repair and Segregation of Chromosomes, Freiburg, Germany, 2010

MORAES, I.C.R.; PINTO-MAGLIO, C.A.F.; LOMBELLO, R.A Cytogenetic characterization and reproductive biology of *Hypericum brasiliense* Choisy. 52º Brazilian Congress of Genetics, Foz do Iguaçu, Brazil, 2006

MORAES, I.C.R.; CORNÉLIO, M.L.; YOUNG, M.C.M.; FACCIO, I.A.P. Estudo da composição química dos óleos voláteis das cascas do fruto de *Punica granatum* L. (Punicaceae) ("Chemistry composition of essencial oils of *Punica granatum* fruits"). 2º Simpósio Brasileiro de Óleos Essenciais, Campinas, Brazil, 2003

PUBLICATIONS

MORAES, I.C.R.; PINTO-MAGLIO, C.A.F.; LOMBELLO, R.A. (2009) "Reproductive biology and cytology of *Hypericum brasiliense* Choisy (Clusiaceae)". Revista Brasileira de Botânica 32(3), 539-544

MORAES, I.; LERMONTOVA, I.; SCHUBERT, I. (2011) "Recognition of *A. thaliana* centromeres by heterologous CENH3 requires high similarity to the endogenous protein". Plant Molecular Biology, in press

Declaration

Herewith I declare that I independently wrote the following doctoral thesis using no other than the sources those which are listed. The principles “Verantwortung in der Wissenschaft” (Responsibility in Science), recommended by the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, were observed.

Date Signature